European University Institute

MAX WEBER PROGRAMME FOR POSTDOCTORAL STUDIES

EUI WORKING PAPERS

MWP 2018/03
Max Weber Programme

The Effect of Job-polarizing Skill Demands on the US Wage Structure

Kory Kantenga

European University Institute
Max Weber Programme

The Effect of Job-polarizing Skill Demands on the US Wage Structure

Kory Kantenga

EUI Working Paper **MWP** 2018/03

ISSN 1830-7728

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

Abstract

I present a quantitative model which accounts for changes in occupational wages, occupational employment shares, and the overall wage distribution. The model reproduces numerous aspects of US cross sectional data observed from 1979 to 2010, notably job and wage polarization. Decompositions reveal changes in production complementarities to be crucial but insufficient to replicate the observed occupational and wage changes. The distribution of worker skills, sorting, and the distribution of skill demands all play pivotal roles. The model indicates skill demands polarized over these three decades, shifting demand away from middle-skilled towards high and – to a lesser extent – low-skilled occupations. I find that industry trends, technological progress, and trade account for up to 57% of changes in skill demands. Information and communications technology spurred demand for jobs requiring interpersonal and social skills in the 1990s. This development appears far more pivotal than the automation of routine jobs concentrated in the manufacturing and construction sectors.

Keywords

Wage Inequality, Job Polarization, Human Capital, Sorting

JEL: D83, D84, J24, J31, J6, O15

I thank Kenneth Burdett, Dirk Krueger, Philipp Kircher, and Iourii Manovskii for invaluable suggestions. Comments and discussions at various seminars at the University of Pennsylvania and the European University Institute were greatly appreciated. All errors are my own.

Kory Kantenga
Max Weber Fellow, 2017-2018, European University Institute
and
Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA, 19104-6297 USA. E-mail: kantenga@sas.upenn.edu

1 Introduction

The US labor market has undergone major changes over the past few decades, affecting what jobs workers do and what their jobs pay. These changes include greater differences in pay (rising wage inequality) and more jobs in high and low-pay occupations versus middle-pay occupations (job polarization). Evidence suggests labor demand-shifting factors ranging from technological change to globalization explain both phenomena. However, changes in occupations and wages sometimes appear counter to these demand-side explanations. For instance, low-paid workers' wages fell relative to the median wage in the 1980s but rose relative to the median in the 1990s. If relative demand shifts drive these wage changes, then we might also expect to see a relative drop in low-paid occupational employment in the 1980s and rise in the 1990s. However, occupational employment changes looked similar in both decades even though wages changes did not.[1] This paper aims to reconcile changes in the occupational structure (average wages and employment) and the wage distribution. In doing so, I estimate a job search model to match these changes from 1979 to 2010. I use the model to infer the underlying shifts in demand which took place and distill which economic forces account for these shifts.

Importantly, we cannot directly observe the skill demands underlying the occupational and wage distributions. We most readily observe equilibrium job allocations and wages. A model of job selection can isolate changes in skill demand and reconcile wage and occupational patterns. For example, Autor and Dorn (2013) show clerical employment correlates negatively with the risk of a machine replacing the worker, however clerical wages correlate positively with this risk. This pattern seemingly contradicts the narrative that labor-saving technology lowered labor demand in automatable occupations, causing their wages and employment to decline. Based on this narrative, we expect clerical employment and wages to correlate negatively with automation risk. However, selection effects can make sense of a positive correlation for wages. If the most productive workers stay in this occupation as demand falls, then average wages may rise. This example shows how a model of job selection might reconcile counterintuitive changes in occupational employment and wages.

The static, competitive Roy (1951) model provides a strong foundation to model job selection but misses out on a rich set of forces which shape wages and the allocation of jobs. In particular, dynamic incentives and labor market search frictions alter the composition and quality of jobs accepted as well as the distribution of earnings across workers even in the same occupation. For example, labor-saving technology increases the risk of job loss for some workers. Employers may increase the wages of workers who anticipate being replaced by a machine to incentivize them to remain at the job despite increasing unemployment risk.[2] The static Roy model misconstrues such wage increases arising from dynamic incentives. Ignoring such forces may lead to different conclusions regarding how skill demand changed and what drove said change. We cannot directly observe such dynamic tradeoffs, but a model can parse their influence. This paper contributes by presenting a quantitative model rich enough to capture these forces, reconciles wage and occupational changes, and yet remains simple enough to estimate with commonly available data.

Several key challenges emerge when considering a model of job selection and wage setting in this context. First, we must specify what mechanisms allocate workers to jobs and determine their pay in equilibrium. Conclusions about how skill demands changed may differ depending on what mechanisms set wages and allocate workers to jobs. Second, occupations pose a severe computational burden, because there are so many of them. The Dictionary of Occupational Titles (DOT) holds over 12,000 occupational titles. Third,

[1] See Lefter and Sand (2011); Mishel, Schmitt, and Shierholz (2013)

[2] E.g. Kredler (2014).

changes to skill demands, skill supply, and productivity remain unobserved. We must make inference about these objects. The model I employ enriches the Roy framework in just enough ways to overcome each of these challenges.

I build on a state-of-the-art model developed by Lise and Postel-Vinay (2016). Workers and employers search and match in the labor market. Those who meet decide whether to form an employer-employee relationship and bargain to determine wages. Workers make decisions today knowing their decision will affect their position in the labor market tomorrow. They possess heterogeneous, multidimensional skills and use their skills to perform tasks of varying complexity in manual and cognitive dimensions. These tasks define occupations. Tasks reduce the dimensional space of occupations, facilitating the inclusion of variety of occupations in structural estimation. Task complexity characterizes the skill level needed to perform a task. Naturally, differences between the worker's skill level and a job's task complexity characterize skill mismatch (i.e. over/under-qualification). This concept provides a natural framework to analyze job selection. Some workers lack the skill level needed in some dimension for a job but others have it. Wages guide workers away from jobs they perform poorly. The model shares the above features with Lise and Postel-Vinay (2016) but in contrast features changes in productivity and the distribution of skill requirements (i.e. skill demands) over time. We cannot fully observe productivity or the distribution of skills demanded or supplied in the data. However, the model imposes enough structure on the data to allow us to draw inference on these latent objects. I discipline the model's parameters using cross-sectional and longitudinal-based moments from US micro data.

The estimated model reproduces numerous aspects of US cross sectional data observed from 1979 to 2010. These aspects include decadal changes in employment shares and average wages across occupational groups and the rise in wage dispersion. The model also replicates the varying patterns of inequality expansion (1980s, 2000s) and contraction (1990s) at the bottom half of the wage distribution. Given a good model fit, I perform a series of decompositions to dissect how the model reconciles occupational and wage changes. The model sheds light on what circumstances led to wage polarization in the 1990s despite consistent job polarization. Job polarization refers to a rise in the employment shares of low and high-skilled occupations at the expensive of medium-skilled occupations. Wage polarization refers to wage compression in the wage distribution below the median and expansion above it. The model infers that production technology shifted away from general skills to specific skills (e.g. cognitive, manual) in the 1980s and then away from manual skills towards cognitive skills in the 1990s, causing wage polarization during this period. Throughout, the distribution of skill demands shifts from manually complex to cognitively complex tasks, causing job polarization. Changes in the distribution of skill endowments and skill demands are just as important as changes in productivity to match occupational and wage changes. Selection effects play a major role in replicating the data. Estimates of changes in skill demands noticeably differ depending on the agents' horizon of foresight over future changes in productivity and skill demands.

The model permits comparisons of prominent explanations for skill demand shifts and leads to insight about what forces drove changes in skill demand over the 1980s, 1990s, and 2000s. The literature on wage inequality and job polarization propose an array of explanations behind shifts in labor demand ranging from the adoption of labor-saving technology (automation) to increased access to cheap labor abroad (offshoring).[3] Classical, closed economy determinants of labor demand include technology (neutral vs. labor-augmenting vs. capital-augmenting) and the price of physical capital. Open economy considerations include the relative price of import goods, i.e. import competition. I take these prominent factors and examine how well they

[3] This literature is too large to survey here. Autor (2015) extensively surveys technological change.

account for changes in the distribution of skill demands.

Two major issues arise in attempting to evaluate these explanations. First, we often observe either limited or low frequency data regarding technological change. For example, we observe different types of capital adoption (e.g. machinery versus transport equipment) annually in cases where we observe them over a long horizon (10+ years). Low range or low frequency time series data make distinguishing between genuine and spurious correlations challenging. I exploit cross-sectional variation in task complexity between occupations and variation in industry concentration across occupations to overcome this challenge. Second, it remains unclear in many studies how much the factors explored contribute to changes in the occupational and wage distributions at the national level. These studies typically exploit cross sectional variation at the country, industry, firm and local area levels.[4] Studies exploiting variation at the cross-country level do not speak to particular national experiences. Aggregating effects across industries, firms, or local areas is non-trivial, because demand-shifting factors can induce broad, national-level general equilibrium responses like labor reallocation across sectors or spending multiplier effects.[5] These effects may amplify or dampen the overall demand impact of any given factor, which summing up local effects may fail to capture.[6] I take an agnostic stance on how skill demands shifted and use the model to estimate them. The model presents a picture of what happened to skill demands nationally as it represents the whole of the US labor market from 1979 to 2010, circumventing this second challenge.

I perform variance decompositions to measure the contribution of the prominent explanations put forward to explain changes in the distribution of skill demands. First, I consider measures of task content to examine what job characteristics not modeled account for skill demand shifts. Task content differs from task complexity. For example, sales and craftsmen jobs require a medium level of cognitive skills even though the content of each job differs greatly. I map cognitive and manual task complexity in the model to occupational task content in the data. I find that demand increased mainly in task areas that require interpersonal skills like negotiation and persuasion. Meanwhile, demand decreased mainly in areas populated with automatable (i.e. "routine") tasks. However, this risk of automation has little explanatory power after controlling for manufacturing and construction industry trends. In contrast, demand growth in interpersonal task areas remains a large explanatory factor for demand shifts even after controlling for industry trends. Jobs more vulnerable to being shipped overseas (i.e. offshored) actually increased in demand on average in the 1990s, all else equal. These jobs include ones which require high cognitive skills but little face-to-face contact like economists and accountants. Next, I measure the contribution of industry trends, capital adoption (to capture technological change), and import competition from China. Decompositions show information and communications technology (ICT) drove changes in skill demands in the 1990s to a large extent. This evidence supports the narrative that ICT developments spurred much demand for jobs requiring interpersonal and social skills. Machinery and transport equipment adoption provide some explanatory power for changes in the 1980s, while drivers of demand in the 2000s remain more mixed. Overall, industry trends and technological progress explain much of the shifts in skill demands, yet a sizable portion (43%) of these changes remain unexplained.

[4]Firpo, Fortin, and Lemieux (2011) provide a notable exception, however they only look at the impact of some factors on the overall wage distribution − not the occupational employment or occupational wages.

[5]Reallocation refers to labor moving to other area in response to a negative shock. Multiplier effects refer to Keynesian-type spending multipliers.

[6]Acemoglu, Autor, Dorn, Hanson, and Price (2016) argue reallocation and multiplier effects on labor demand caused by increased Chinese import competition take place mostly within local areas and thus aggregating local area effects reflects the national impact.

1.1 Connected Literature

This paper relates to several dense and interconnected literatures. The challenges mentioned provide an organizing principal to parse this dense literature and place this paper into context.

Determining the endogenous allocation of workers to jobs dates back to Roy (1951) whose model remains widely used to frame the endogenous allocation of workers to jobs (Boehm, 2017; Autor and Dorn, 2013; Autor and Handel, 2013; Yamaguchi, 2012). In a standard Roy model, workers possess specific heterogeneous skills, and competitive skill prices allocate workers across jobs. While a good foundation, this setup ignores dynamic decision making and labor market imperfections. Consequently, the Roy model mischaracterizes a set of rich and potentially important outcomes surrounding occupational choice which the dynamic, structural model here captures. Work dating back to Willis and Rosen (1979) supports the notion that workers make dynamic career decisions, forecasting their future earnings to make schooling and occupational choices.[7] Dynamic decisions change selection and wage setting incentives in the presence specific human capital. For example, Chari and Hopenhayn (1991) show wages in declining "vintages" (e.g. manufacturing jobs) face countervailing pressures in the presence of specific human capital. On one hand, new hires need an incentive to acquire and maintain specific skills in a job which has increasingly less productive value elsewhere and high risk of layoff. This pressures employers to pay higher starting wages to fill vacancies. On the other hand, older workers become stuck in this occupation over time, so the employer need not compensate them as much to stay.[8] Average wages in the occupation may rise if the former force dominates. The Roy model can only reconcile wage increases with higher demand or selection of better workers into the job. Neither need occur in this example. The interaction between labor market imperfections and dynamic considerations (i.e. the inability to easily move to a new job) drive wage dynamics here. Thus, the competitive lens of the Roy model misinterprets this scenario.

Task-specific, heterogeneous human capital provides a way to incorporate occupations without a heavy computational burden and model unobserved skill evolution. Task-specific means this human capital only helps perform a specific task. Workers differ in their stock of human capital, making it heterogeneous. Task-specific human capital emerged in the job/wage polarization literature to explain non-monotone changes across the wage distribution (Acemoglu and Autor, 2011). Papers apply this framework to understand phenomena like occupational mobility (Sanders, 2016) and why skills reward differently across occupations (Yamaguchi, 2012).[9] Even more papers use it in the Roy occupational choice framework to explain the drivers of job and wage polarization. Datasets like The Dictionary of Occupational Titles (DOT) and O*NET provide information on tasks to estimate this class of models, while datasets like the National Longitudinal Survey of Youth (NLSY) provides information to estimate pre-labor market entry skills which evolve according to the model. The model here adopts this framework.

Structural and reduced form literatures provide differing ways to deal with unobserved skill demand and supply. In the structural literature, Lindenlaub (2017) estimates the static, competitive equilibrium of a multidimensional assignment model. Task, skill, and wage data identify skill supply and demand as well as productivity in the model. She then uses the model to explain wage changes over the 1990s and 2000s. Changes in production technology parameters come as an unanticipated shock and wages fully adjust. The model provides valuable insights into the mechanics of wage polarization where workers may

[7] E.g. Keane and Wolpin (1997); Heckman, Lochner, and Taber (1998). Dynamic decisions refers to the feature that workers take the future into account for their decision today. It seems unlikely workers suddenly become myopic entering the labor market.

[8] Their skills are specific and increasingly less valued elsewhere.

[9] Sanders and Taber (2012) provide an extensive overview of this literature.

switch ranks. While useful to examine wage changes, the assignment equilibrium of Lindenlaub (2017) remains unsuitable to address job polarization or other changes in the occupational structure.[10] Changes in production complementarities drive wage polarization in this frictionless, static assignment model. However, this frictional, dynamic model says changes in the distribution of skill supply and skill demand appear just as important as changes in productivity when matching both occupational and wage changes.

I build directly on Lise and Postel-Vinay (2016). Their multidimensional skill, search model specifies how skills evolve to gain insight on the role of skill accumulation and mismatch over a worker's life cycle. There, changes in the occupation structure only arise from job selection − not changes in the distribution of skill requirements or productivity. The distribution of skill requirements and productivity parameters remain fixed.[11] They estimate their model to match moments in the NLSY1979 male cohort and examine the lose due to skill mismatch. In contrast, I examine a transition path where the distribution of skill requirements (i.e. demands) and productivity evolve and match moments on both the NLSY1979 cohort (to discipline model parameters) and the cross-sectional distributions of wages and occupations over time.

The reduced form literature uses econometric techniques to identify demand shifts consistent with wage and occupational changes. It takes equilibrium employment and wage outcomes as given and aims to separate out selection effects without estimating the underlying structural model.[12] This literature proposes a variety of demand shifting factors like import penetration and metrics for technology adoption to pin down demand shifts. Papers exploit time-series, cross-sectional variation in these factors across firms (Bresnahan, Brynjolfsson, and Hitt, 2002; Bartel, Ichniowski, and Shaw, 2007), industries (Autor, Katz, and Krueger, 1998), countries (Michaels, Natraj, and Van Reenen, 2014; Goos, Manning, and Salomons, 2014), and local areas (Autor, Dorn, and Hanson, 2013; Autor and Dorn, 2013; Acemoglu and Restrepo, 2017). I view my structural approach as complementary to this literature. It provides a comprehensive framework in which to compare and contrast some of the demand-shifting economic forces that these papers present. The model's job selection mechanism makes the effect of changes in skill demands on wages ambiguous at both the individual and occupational levels. This ambiguity allows the model to match seemingly contradictory changes in employment and wages. Such ambiguity also exists in competitive models on the aggregate level like Boehm (2017) and Kredler (2014). Boehm (2017) employs a static, perfectly competitive Roy model where selection effects (also referred to as sorting) generate ambiguity at the aggregate level. Increased demand for a task reallocates workers across jobs. Occupational wages may rise or fall depending on the skill distribution of the workers who move. In this paper, workers reallocate, but search and matching frictions affect reallocation. Kredler (2014) employs a dynamic model of human capital and technological change. Wage dynamics generate ambiguity at the occupational level.[13] Wages rise as an occupation contracts to compensate entering workers for a shorter career. Similarly, wages may rise as prospects of a job-to-job move worsen in my environment. Experienced workers face wage loses as their skills become obsolete.[14] All else equal, average wages within an occupation may rise or fall with occupational contraction, depending on the

[10] See Footnote 57 in Lindenlaub (2017).

[11] An obvious way around this drawback is to estimate the equilibrium of the Lise and Postel-Vinay (2016) over sub-periods where skill requirements and technology remain fixed. I show this estimation approach matches wages changes well but fails to match wage and employment changes at the occupational level. It also precludes any analysis on the role of changing expectations in job polarization and wage determination

[12] In contrast, the structural approach puts structure on selection effects and equilibrium outcomes, aiming to use the model to make inference about the data.

[13] Kredler (2014) constructs a dynamic model based on the two-period model of Chari and Hopenhayn (1991).

[14] Their skills become obsolete, because human capital is specific to their occupation or vintage. Entering workers do not possess as much occupation-specific human capital, so obsolescence does not drive down starting wages. In Kredler (2014), human capital solely depends on the level of experience in an occupation or vintage whereas here it can also depend on work history.

distribution of entrants and experienced workers.

2 Model

2.1 The Environment

Time is discrete. The economy consists of workers and jobs. Workers enter and exit the labor market exogenously. Workers may be employed, unemployed, or out of the labor force. They live finite lives and possess human capital also referred to as skills. All workers possess a non-separable bundle of general, cognitive and manual skills denoted by $\mathbf{x} \in \mathcal{X}$. Workers use their skills to do tasks. Cognitive and manual skills are task-specific, meaning manual skills do not contribute to doing cognitive tasks and vice versa. General skills affect the overall efficiency level doing any task. Skills reflect task complexity.[15]

Employers or firms offer a job as in the standard Mortensen-Pissarides framework (Mortensen and Pissarides, 1999). A job consists of a non-separable bundle of cognitive and manual skill requirements (or demands) denoted by $\mathbf{y} \in \mathcal{Y}$. Skill requirements differ according to the firm's production technology. Workers search for jobs and supply their skills to a firm with whom they match. Skill requirements reflect the task complexity required for a job. Employers post job vacancies and draw skill requirements from the distribution $\mathcal{F}(\mathbf{y})$. Matched employers use their technology and the worker's skills to produce output. Employers pay workers wages thereby splitting the total value (or surplus) created from the worker-employer match. $f(\mathbf{x}, \mathbf{y})$ is the flow value of output from a match between a worker with skills $\mathbf{x}$ and a firm requiring skills $\mathbf{y}$ where $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$. $c(\mathbf{x}, \mathbf{y})$ is the flow disutility of labor for worker $\mathbf{x}$ at firm $\mathbf{y}$ where $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$. $b(\mathbf{x})$ is the flow utility of an unemployed worker $\mathbf{x}$ where $b : \mathcal{X} \to \mathbb{R}_+$. In what follows, the subscript of t denotes that the function is time dependent.

Workers and firms have a common discount factor $\tilde{\beta}$. As mentioned, worker transitions in and out of the labor force are exogenous. Workers entering the labor market at time t draw their skills from an exogenous distribution $\mathcal{V}_t(\mathbf{x})$. Workers enter the labor market at time t, aged a and draw initial skills denoted by $\mathbf{x}(0)$. They exit the labor market permanently with age-dependent probability ξ_a and exit with certainty at age 65.[16] The distribution of worker skills in the economy is denoted by $\mathcal{W}_t(\mathbf{x})$. The skill requirements distribution, $\mathcal{F}_t(\mathbf{y})$, evolves exogenously whereas the worker skills distribution, $\mathcal{W}_t(\mathbf{x})$, evolves endogenously when there is human capital evolution as I will describe shortly.

Workers and firms engage in random search in a single labor market. Employed and unemployed workers encounter an employer in each period with probabilities $\mathbb{M}_{e,t}$ and $\mathbb{M}_{u,t}$, respectively. Given an encounter, a job offer is drawn from the commonly known distribution $\mathcal{F}(\mathbf{y})$. Jobs may be destroyed with exogenous probability δ. Enduring matches face a permanent productivity shock with probability ω where the firm

[15] Task-specific skills are coarser and more transferrable than occupation-specific skills. This task-specific framework based on task complexity has two important advantages. First, the framework accomodates many occupations with a much smaller number of parameters. With occupation-specific skills, the number of model parameters (e.g. productivity levels) increases with each additional occupation. In contrast, this number does not grow with the number of occupations with task-specific skills, allowing us to accomodate many occupations. Second, this framework provides a natural explanation for why different occupations have similar pay. Similar pay is due to the similar complexity of the tasks these occupations require (Yamaguchi, 2012).

[16] In Cortes, Jaimovich, and Siu (2016), workers decide whether to enter the labor market before deciding where to work and do not know their skill level ex-ante. They decide based the realization of a stochastic process and the expected returns to working. Implicitly here, the worker entry-exit decision depends on age, a stochastic process, and mandatory retirement at 65. I use the corresponding reduced-form probabilities in this model. Explicitly, worker entry-exit may depend on factors like the value of home production or leisure. Exogenous entry-exit probabilities will capture these decisions so long as the ex-ante expected labor market return depends only on age. This will be the case if workers learn their specific skills only after entry.

draws new skill requirements from $\mathcal{F}(\mathbf{y})$.[17] Workers and firms take the distribution of skill requirements as given at time t and forecast it over future dates.[18] A worker's task-specific skills at a job requiring skills $\mathbf{y}$ evolve according the law of motion $h : \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$ where

$$\mathbf{x}(t+1) = h(\mathbf{x}(t), \mathbf{y}), \tag{1}$$

thus (1) defines human capital (or skill) evolution at job $\mathbf{y}$. I assume h satisfies the following:

$$\lim_{t\to\infty} h(\mathbf{x}(t), \mathbf{y}) = \mathbf{y}, \tag{2}$$

$$\lim_{t\to\infty} h(\mathbf{x}(t), \mathbf{0}) = \underline{\mathbf{x}}. \tag{3}$$

(3) says the skills for unemployed workers (for whom $\mathbf{y} = \mathbf{0}$) depreciate towards a lower bound ($\underline{\mathbf{x}}$) in the support $\mathcal{X}$ as the duration of their unemployment spell grows. (2) defines learning-by-doing. A worker's skills converge to the skill requirements of the job as they spend time on-the-job. Workers with skills exceeding those required lose their excess skill level over time, while workers learn skills on-the-job for which they remain deficient. I call an employed worker over-qualified in a skill dimension when that worker's skill level exceeds the required skill level and under-qualified in a skill dimension when that skill level falls short of the required skill for the job. With learning-by-doing, job selection determines human capital evolution, because the job selected determines gains and loses of skill.[19] Thus, the distribution of worker skills becomes endogenous if $h(\mathbf{x}(t), \mathbf{y}) \neq \mathbf{x}\ \forall(\mathbf{x}, \mathbf{y}, t)$.

2.1.1 Timing

At the start of each period, the worker is employed, unemployed, or out of the labor force and their skills have evolved accordingly. For an employed worker, the match breaks up exogenously with probability δ and the worker leaves the work force with probability ξ_a. If still employed, then worker produces $f(\mathbf{x}, \mathbf{y})$ with their current firm. Next, the worker meets a new employer with probability $\mathbb{M}_{e,t}$ and then receives a job offer ($\mathbf{y}$). If a meeting occurs, the worker and potential employer decide whether to form the match and then proceed to negotiate the split of the surplus. If they both accept the match, then the employed worker starts the next period with the new employer, leaving the current employer. If the worker does not meet an employer, then the current match may experience a permanent shock to skill requirements (w.p. ω). If faced with the permanent shock, the employer and worker decide whether to remain matched or separate.[20] The worker starts the next period unemployed in the case of a separation following the shock.

An unemployed worker ($\mathbf{x}$) receives an exogenous utility flow $b(\mathbf{x})$ at the start of the period. Next, the worker meets an employer with probability $\mathbb{M}_{u,t}$ and then receives a job offer ($\mathbf{y}$). If a meeting occurs, the employer and worker decide whether to form the match and proceed to negotiate the split of the surplus. If they both accept the match, then the newly employed worker starts the next period with the employer, barring a separation or labor market exit at the start of the next period. If the match does not form or no meeting takes place, then the worker stays unemployed the next period, barring a labor market exit at the start of the next period. Workers out of the labor force exogenously enter as unemployed at the start of the

[17] The distribution of firm skill demands evolves exogenously. Although unmodelled, these skill requirements evolve with technological change. I introduce technological innovation on-the-job through this permanent shock to skill requirements.

[18] I consider the cases where all agents have no foresight and perfect foresight over this distribution.

[19] Job selection is equivalent to task or occupational selection in this model, because the combination of tasks a worker selects defines their occupation.

[20] Matches terminate mutually if the surplus falls below zero, so workers quitting is equivalent to employers firing them in this model. The worker may quit to go into unemployment in order to search again with the meeting rate $\mathbb{M}_{u,t}$.

period.

At the start of each period, an unmatched employer posts a vacancy at cost τ and meets a worker with probability $\mathbb{M}_{v,t}$. Upon meeting a worker, the employer draws skill requirements ($\mathbf{y}$) and then decides whether to form a match with the worker ($\mathbf{x}$). If the match forms, then they negotiate the split of the surplus and begin producing together the next period. Matched employers produce $f(\mathbf{x}, \mathbf{y})$ with the worker at the start of each period and then engage in negotiations if the worker meets another employer who makes a poaching offer. Matched employers whose workers do not meet another employer may experience a permanent shock to skill requirements (w.p. ω). If faced with the permanent shock, the employer and worker decide whether to remain matched or separate. Newly unmatched employers may post a vacancy tomorrow in the way described or freely exit the labor market. Employers outside the labor market may freely enter at the start of the period as an unmatched employer.

2.1.2 Bargaining Protocol

Workers and employers bargain over the total value (or surplus) generated by the match. The outcome of this bargaining process determines the split of the surplus. The bargaining protocol follows the sequential auction model of Cahuc, Postel-Vinay, and Robin (2006). Unemployed workers with bargaining power $\lambda \in [0, 1]$ bargain with employers à la Nash. Hence, unemployed workers take a share of the surplus equal to λ. Employers attempting to poach employed workers compete with the worker's current employer. If an employed worker meets an employer offering skill requirements ($\mathbf{y}'$), then the two employers engage in Bertrand competition over the share of the surplus to give the worker. As result, the worker receives a value equal to at least the surplus of the employer with whom the worker generates lower surplus. This value is the worker's outside option in the bargaining process. The worker and employer with higher surplus then engage in Nash bargaining over the surplus amount exceeding the worker's outside option. Thus, a job-to-job transition only occurs when the surplus for the poaching employer exceeds that of the current employer.

To illustrate the process, let $S(\mathbf{x}, \mathbf{y})$ denote the surplus of a match of employer with skill requirements $\mathbf{y}$ and worker with skills $\mathbf{x}$. Let $W(\mathbf{x}, \mathbf{y}, \sigma)$ denote the value the worker receives in the match and σ denote the share of the surplus received. Suppose a meeting on-the-job occurs and $S(\mathbf{x}, \mathbf{y}') \geq S(\mathbf{x}, \mathbf{y}) > W(\mathbf{x}, \mathbf{y}, \sigma)$ so that the poaching employer with $\mathbf{y}'$ generates higher surplus with the worker than the current employer. The offer of $\mathbf{y}'$ triggers a bidding war between the two employers, because the worker expects to gain from renegotiating the wage contract, $W(\mathbf{x}, \mathbf{y}, \sigma)$. The worker stands to gain in the case where $S(\mathbf{x}, \mathbf{y}) > W(\mathbf{x}, \mathbf{y}, \sigma)$. Bertrand competition causes employers $\mathbf{y}$ and $\mathbf{y}'$ to bid until $W = S(\mathbf{x}, \mathbf{y})$ at which point employer $\mathbf{y}$ loses the bidding war. Then, employer $\mathbf{y}'$ and the worker Nash bargain over $S(\mathbf{x}, \mathbf{y}') - S(\mathbf{x}, \mathbf{y})$ where the worker has bargaining power λ. Hence, the share of the surplus for the worker at the new employer (σ') is

$$\sigma' = \sigma(\mathbf{x}, \mathbf{y}', \mathbf{y}) = \frac{\lambda[S(\mathbf{x}, \mathbf{y}') - S(\mathbf{x}, \mathbf{y})] + S(\mathbf{x}, \mathbf{y})}{S(\mathbf{x}, \mathbf{y}')} = \lambda + (1 - \lambda)\frac{S(\mathbf{x}, \mathbf{y})}{S(\mathbf{x}, \mathbf{y}')} \in (0, 1]. \tag{4}$$

The employed worker takes a value $W(\mathbf{x}, \mathbf{y}', \sigma')$ equal to the lower surplus of the two employers plus a share of the surplus gain from the job-to-job move.[21] The corresponding worker's surplus share consists of the unemployed worker's share (λ) and an additional amount generated by competition between the employers for the worker.[22]

[21] I assume the share of the surplus stays constant until an on-the-job meeting triggers renegotiation. Assuming the share stays constant until renegotiation does not affect mobility decisions but does affect the time profile of wage payments as Lise and Postel-Vinay (2016) note. Total value (or surplus) determines mobility. I also assume unemployed workers accept job offers when indifferent.

[22] I elaborate on the reasons for using this protocol in Appendix A.2.

2.2 Worker's Problem

Let z_t denote the aggregate state variables $\mathcal{F}_t$, f_t, $\mathbb{M}_{e,t}$, and $\mathbb{M}_{u,t}$. Let any function $T(\cdot; z_t)$ be denoted by $T_t(\cdot)$ and $\mathbb{E}_t$ denote the expectation over z_{t+1}.[23] As mentioned, $\mathbf{y}$ denotes the skill requirements of the current employer of a worker. $\mathbf{y}$ consists of cognitive (y_c) and manual (y_m) skill requirements. $\mathbf{x}$ denotes the skills of the worker which evolve to $\mathbf{x}'$ next period. $\mathbf{x}$ consists of cognitive skills (x_c), manual skills (x_m), general skills (x_g), and age (a). For workers, I define an age effective discount factor $\beta_a = \tilde{\beta}(1 - \xi_a)$. Denote the value functions for an unemployed and employed workers at time t as $U_t(\mathbf{x})$ and $W_t(\mathbf{x}, \mathbf{y}, \sigma)$, respectively. σ denotes the employed workers endogenous share of the total surplus. I denote the total surplus at time t by $S_t(\mathbf{x}, \mathbf{y})$. I assume σ remains constant prior to renegotiation.[24] Since the worker receives a constant share σ of the surplus $S_t(\mathbf{x}, \mathbf{y})$, $W_t(\mathbf{x}, \mathbf{y}, \sigma) = \sigma S_t(\mathbf{x}, \mathbf{y}) + U_t(\mathbf{x})$. I assume linear utility in wage income.[25] An unemployed worker receives a flow income $b(\mathbf{x})$. Thus, the unemployed worker's value function $U_t(\mathbf{x})$ imposing the bargaining protocol solves

$$U_t(\mathbf{x}) = b(\mathbf{x}) + \beta_a \mathbb{E}_t U_{t+1}(\mathbf{x}') + \beta_a(1 - \delta)\lambda \mathbb{M}_{u,t} \mathbb{E}_t \int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}', \mathbf{y})\} \, \mathrm{d}\mathcal{F}_t(\mathbf{y}). \tag{5}$$

where $\mathbf{x}' = h(\mathbf{x}, \mathbf{0})$ for x_m and x_c. The value of being unemployed consists of the flow income $b(\mathbf{x})$, the age-discounted present value of being unemployed tomorrow, and the present value of the expected share of the surplus if the worker finds employment.

Let $w_t(\mathbf{x}, \mathbf{y}, \sigma)$ be the wage implementing the employed worker's wage contract at time t. The employed worker's value function $W_t(\mathbf{x}, \mathbf{y}, \sigma)$ given σ and imposing the bargaining protocol solves

$$\begin{aligned} W_t(\mathbf{x}, \mathbf{y}, \sigma) \quad &= w_t(\mathbf{x}, \mathbf{y}, \sigma) - c(\mathbf{x}, \mathbf{y}) + \beta_a \mathbb{E}_t U_{t+1}(\mathbf{x}') + \beta_a(1 - \delta)(1 - \mathbb{M}_{e,t})\sigma \mathbb{E}_t \tilde{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \\ &\qquad \beta_a(1 - \delta)\mathbb{M}_{e,t} \times \\ &\mathbb{E}_t \int_{\mathcal{Y}} \max\{\sigma \widehat{S}_{t+1}(\mathbf{x}', \mathbf{y}), \widehat{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \lambda[S_{t+1}(\mathbf{x}', \mathbf{y}') - \widehat{S}_{t+1}(\mathbf{x}', \mathbf{y})]\} \mathrm{d}\mathcal{F}_t(\mathbf{y}'), \end{aligned} \tag{6}$$

where

$$\begin{aligned} \widehat{S}_{t+1}(\mathbf{x}, \mathbf{y}) &= \max\{S_{t+1}(\mathbf{x}, \mathbf{y}), 0\} \\ \tilde{S}_{t+1}(\mathbf{x}', \mathbf{y}) &= (1 - \omega)\widehat{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \omega \int_{\mathcal{Y}} \max\{S_{t+1}(\mathbf{x}', \mathbf{y}'), 0\} \mathrm{d}\mathcal{F}_t(\mathbf{y}') \end{aligned}$$

subject to (1). As in Hagedorn, Law, and Manovskii (2017), I assume a small offer writing cost ϵ prevents employers with lower surplus than the current employer from engaging in Bertrand competition if an on-the-job meeting occurs so that $\sigma' = \sigma$ in equilibrium if $S(\mathbf{x}, \mathbf{y}) \geq S(\mathbf{x}, \mathbf{y}') > W(\mathbf{x}, \mathbf{y}, \sigma)$.[26] The value of employment with firm $\mathbf{y}$ consists of the wage less disutility of labor, the present value of unemployment, and the share of the surplus if the worker does or does not meet a firm while searching on-the-job or experiences a permanent change to skill requirements after no meeting on-the-job.

[23] Thus, any function subscripted with t also has the argument z_t.

[24] This assumption pins down wages in the model, because wages adjust to deliver this constant surplus split.

[25] I thereby assume risk neutrality for workers and firms. The assumption significantly increases the tractability of the model at the cost of precluding any kind of meaningful welfare analysis.

[26] This restriction prevents bidding up of wages on-the-job in order to restrict attention to human capital in terms of producing wage growth over job tenure in the model. An obvious extension would be to allow both human capital accumulation and bidding up of the share of surplus on-the-job (so-called job shopping). In Appendix A.3, I show the wage and surplus function without this restriction.

2.3 Employer's Problem

As described, an unmatched employer decides whether to post a vacancy at time t and then draws skill requirements from $\mathcal{F}_t(\mathbf{y})$ if it meets a worker. Employers draw a new $\mathbf{y}$ at each worker meeting. Hence, the value of a vacancy is the same for all unmatched employers ex-ante. Let τ_t be the cost of posting the vacancy at time t and let V_t be the value of this vacancy posting. Let $P_t(\mathbf{x}, \mathbf{y}, \sigma)$ be the value of producing with a worker of type $\mathbf{x}$ and delivering surplus share σ. Let $\mathbb{C}_{e,t}$ be the probability of meeting an employed worker and $\mathbb{C}_{u,t}$ be the probability of meeting an unemployed worker − all conditional on meeting a worker. Then the value of a vacancy V_t solves

$$\begin{aligned} V_t \quad &= -\tau_t + (1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{u,t}(1-\lambda)\mathbb{E}_t \int\limits_{\mathcal{Y}} \int\limits_{\mathcal{X}|u} \beta_a \max\{0, S_{t+1}(\mathbf{x}, \mathbf{y})\} \mathrm{d}\mathcal{F}_t(\mathbf{y}) \mathrm{d}\mathcal{W}_t(\mathbf{x}|u) + \\ &\qquad (1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{e,t}(1-\lambda) \times \\ &\mathbb{E}_t \int\limits_{\mathcal{Y}} \int\limits_{\mathcal{Y}\times\mathcal{X}|e} \beta_a \max\{0, S_{t+1}(\mathbf{x}, \mathbf{y}) - \widehat{S}_{t+1}(\mathbf{x}, \mathbf{y}')\} \mathrm{d}\mathcal{F}_t(\mathbf{y}) \mathrm{d}\mathcal{W}_t(\mathbf{x}, \mathbf{y}'|e) \end{aligned} \tag{7}$$

where $\mathcal{W}_t(\mathbf{x}|u)$ and $\mathcal{W}_t(\mathbf{x}, \mathbf{y}'|e)$ are the distributions of unemployed workers and employer-employee matches at time t, respectively. I assume free entry of employers which drives the value of vacancy to zero so that

$$\begin{aligned} \tau_t = \quad &(1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{u,t}(1-\lambda)\mathbb{E}_t \int\limits_{\mathcal{Y}} \int\limits_{\mathcal{X}|u} \beta_a \max\{0, S_{t+1}(\mathbf{x}, \mathbf{y})\} \mathrm{d}\mathcal{F}_t(\mathbf{y}) \mathrm{d}\mathcal{W}_t(\mathbf{x}|u) + \\ &\qquad (1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{e,t}(1-\lambda) \times \\ &\mathbb{E}_t \int\limits_{\mathcal{Y}} \int\limits_{\mathcal{Y}\times\mathcal{X}|e} \beta_a \max\{0, S_{t+1}(\mathbf{x}, \mathbf{y}) - \widehat{S}_{t+1}(\mathbf{x}, \mathbf{y}')\} \mathrm{d}\mathcal{F}_t(\mathbf{y}) \mathrm{d}\mathcal{W}_t(\mathbf{x}, \mathbf{y}'|e). \end{aligned} \tag{8}$$

Alternative assumptions on free entry and the timing when employers learn their types are possible[27], however I choose this timing and the free entry assumption for tractability as in Lise and Postel-Vinay (2016).

The value of producing solves

$$\begin{aligned} P_t(\mathbf{x}, \mathbf{y}, \sigma) \quad &= f_t(\mathbf{x}, \mathbf{y}) - w_t(\mathbf{x}, \mathbf{y}, \sigma) + \beta_a(1-\delta)(1-\mathbb{M}_{e,t})(1-\sigma)\mathbb{E}_t\tilde{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \\ &\qquad \beta(1-\delta)\mathbb{M}_{e,t}\mathbb{E}_t\big[\max\{0, (1-\sigma)S_{t+1}(\mathbf{x}', \mathbf{y})\} \cdot \rho(\mathbf{x}, \mathbf{y})\big] \end{aligned} \tag{9}$$

where

$$\rho(\mathbf{x}, \mathbf{y}) = \int\limits_{\mathcal{Y}} \mathbb{1}\{S_{t+1}(\mathbf{x}', \tilde{\mathbf{y}}) < S_{t+1}(\mathbf{x}', \mathbf{y})\} \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})$$

subject to (1). $\rho(\mathbf{x}, \mathbf{y})$ is the probability the worker at $\mathbf{y}$ does not draw an employer with higher surplus.[28] The matched employer receives output less wages and the share of the surplus from producing next period which depends on whether or not another employer poaches the worker. If the worker does not meet another employer, then the current employer draws new skill requirements (w.p. ω).[29] $P_t(\mathbf{x}, \mathbf{y}, \sigma) = (1-\sigma)S_t(\mathbf{x}, \mathbf{y}) + V_t$ since the employer takes a constant share of the surplus,. It follows that the total surplus of a match $(\mathbf{x}, \mathbf{y})$ is

$$S_t(\mathbf{x}, \mathbf{y}) = \underbrace{W_t(\mathbf{x}, \mathbf{y}, \sigma) - U_t(\mathbf{x})}_{\sigma S_t(\mathbf{x}, \mathbf{y})} + \underbrace{P_t(\mathbf{x}, \mathbf{y}, \sigma) - V_t}_{(1-\sigma)S_t(\mathbf{x}, \mathbf{y})}. \tag{10}$$

[27] See Hagedorn, Law, and Manovskii (2017). However, their model is one dimensional, and the distribution of $\mathbf{y}$ is normalized to uniform for identification. This model does not impose these restrictions.

[28] $\mathbb{1}\{\cdot\}$ denotes the indicator function.

[29] The match only draws new skill requirements if the worker does not meet another employer. I impose this structure to make the model more tractable in terms of solving for the surplus.

2.4 Surplus, Wages, and Equilibrium Concept

Now we can derive the surplus function using (5), (6), (9), and the free entry assumption which implies that V_t equals zero. Given meet probabilities, we can also solve $S_t(\mathbf{x},\mathbf{y})$ backwards, because $\beta_a = 0$ ($\xi_a = 1$) for workers aged 65 and older. These workers leave the labor force due to mandatory retirement as stated earlier. The surplus for a match where the worker retires next period is

$$S_t(\mathbf{x},\mathbf{y}) = f_t(\mathbf{x},\mathbf{y}) - c(\mathbf{x},\mathbf{y}) - b(\mathbf{x}) \tag{11}$$

which is just the static flow of the surplus. For non-retiring workers, the surplus is

$$\begin{aligned} S_t(\mathbf{x},\mathbf{y}) = \quad & f_t(\mathbf{x},\mathbf{y}) - c(\mathbf{x},\mathbf{y}) - b(\mathbf{x}) + \beta_a(1-\delta)\mathbb{E}_t\Bigg[-\lambda \mathbb{M}_{u,t} \int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\} \, \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}}) \; + \\ & \big(1-\mathbb{M}_{e,t}\big)\tilde{S}_{t+1}(\mathbf{x}',\mathbf{y}) + \mathbb{M}_{e,t} \cdot \rho(\mathbf{x},\mathbf{y}) \cdot \max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\} + \\ & \mathbb{M}_{e,t} \cdot (1-\rho(\mathbf{x},\mathbf{y})) \cdot \Big[\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) + \lambda\big(\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}) - \widehat{S}_{t+1}(\mathbf{x}',\mathbf{y})\big)\Big]\Bigg], \end{aligned} \tag{12}$$

$$\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}) = \frac{\int_{\mathcal{Y}} \mathbb{1}\{\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) < S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\} \cdot S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}}) \, \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}{\int_{\mathcal{Y}} \mathbb{1}\{\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) < S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\} \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}.$$

Assuming the match survives to next period (w.p. $1-\delta$), the surplus consists of the static flow and the continuation value. The continuation value consists of four terms. The first term reflects that the worker can quit and search again as an unemployed worker and expects to obtain the value shown. It enters negatively into the surplus, because the incentive to form the match falls if the worker's incentive to quit the next period rises. The second and third terms consist of two parts. The first part is the probability that the worker does not leave the match. The worker only stays if 1) a meeting does not place (w.p. $1-\mathbb{M}_{e,t}$) or 2) a meeting takes place but the poaching employer draws skill requirements that do not deliver higher surplus (w.p. $\mathbb{M}_{e,t}\rho$) given the worker's $\mathbf{x}$. The second part is the surplus next period, barring a mutual separation due to a negative match surplus. Thus, the second and third terms are the value coming from the expectation to remain in the match. Naturally, the fourth term is the value coming from the expectation to leave the match for another job. $\mathbb{M}_{e,t}(1-\rho(\mathbf{x},\mathbf{y}))$ is the probability of meeting an employer who draws skill requirements that deliver a higher surplus for the worker's $\mathbf{x}$. The second part of this last term consists of the expected value the worker obtains from transitioning to a new employer.

(12) shows how the distribution of skill requirements and other parameters governing the surplus affect the total value of a match and consequently match formation and continuation. The worker and the employer care about who the worker can meet next, because some of the total gain from a job-to-job move goes to the worker. The employer extracts some of that gain today. This potential gain affects the current surplus and in turn affects match formation and match continuation.[30] Thus, expected changes in $\mathcal{F}_t(\mathbf{y})$ due to drivers of job polarization also influence current job selection through the value of a potential or current match.

The effect of changes in $\mathcal{F}_t(\mathbf{y})$ on the value of a match are generally ambiguous. Let us refer to the probability of drawing a better match in terms of surplus $(1-\rho)$ as the worker's job prospects.[31] Suppose

[30] I use match formation and job selection interchangeably. The employer and worker do not care about who the employer can meet next except vis-à-vis the option value of the employer searching again (i.e. the value of a vacancy). Employers do not search for replacement employees on-the-job as workers search for new employers. Employers searching on-the-job to replace the worker adds an additional and potentially interesting layer of complexity that I do not take on here.

[31] Better also refers to cases where surplus is at least as good.

ω is zero and $\mathcal{F}_t$ changes once permanently such that $\rho(\mathbf{x}^*, \mathbf{y}^*)$ rises for the worker $\mathbf{x}^*$ at employer $\mathbf{y}^*$. In other words, the worker $\mathbf{x}^*$'s job prospects worsen at the current employer $\mathbf{y}^*$. This change lowers the option value of searching again as an unemployed worker and increases the value of continuing the match tomorrow. Both of which increase the surplus. Intuitively, the current match becomes more valuable to the worker as job prospects worsen. However, worsening prospects ambiguously affect the expected value from leaving to a better job. It lowers the probability of the worker finding a better match $(1 - \rho)$, but it may increase or decrease the expected value of a new match, $\bar{S}_{t+1}$. This effect depends on how $\mathcal{F}$ changes. Suppose the probability mass on matches with the highest surplus for $\mathbf{x}^*$ move to matches with the lowest surplus, then $\bar{S}_{t+1}$ falls. In this case, the expected value from leaving the current match falls, lowering the value of the current match. This effect offsets the increase from the first three continuation terms. Thus, worsening job prospects have an ambiguous effect on the value of a match. Hence, we cannot determine a priori the selection effects of a change in $\mathcal{F}_t(\mathbf{y})$. Intuitively, this effect should be ambiguous. Worsening job opportunities make a job both more valuable and less valuable. A job becomes more valuable when finding a better one becomes more difficult, but a lack of future opportunities makes the job less valuable. These opposing considerations complicate predicting the allocative impact of job-polarizing skill requirements.

An employer delivers the worker's share of the surplus through wages. Combining (12), $W_t(\mathbf{x}, \mathbf{y}, \sigma) = \sigma S_t(\mathbf{x}, \mathbf{y}) + U_t(\mathbf{x})$, and substituting in (6) produces the following wage equation

$$\begin{aligned} w_t(\mathbf{x}, \mathbf{y}, \sigma) = \quad & \sigma f_t(\mathbf{x}, \mathbf{y}) + (1-\sigma)c(\mathbf{x}, \mathbf{y}) + (1-\sigma)b(\mathbf{x}) + (1-\sigma)\beta_a(1-\delta) \times \\ & \mathbb{E}_t\Bigg[\lambda \mathbb{M}_{u,t} \int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}', \mathbf{y})\} \, \mathrm{d}\mathcal{F}_t(\mathbf{y}) - \\ & \mathbb{M}_{e,t} \cdot (1 - \rho(\mathbf{x}, \mathbf{y}))\Big(\lambda \bar{S}_{t+1}(\mathbf{x}', \mathbf{y}) + (1-\lambda)\widehat{S}_{t+1}(\mathbf{x}', \mathbf{y})\Big)\Bigg]. \end{aligned} \tag{13}$$

The first three terms consist of the worker's share of the static surplus (11) plus the labor disutility, $c(\mathbf{x}, \mathbf{y})$, and outside option, $b(\mathbf{x})$, flows. The potential gains from a transition to unemployment and a transition to another employer make up the continuation value's wage contribution. The wage increases with the attractiveness of unemployment in order to deliver the share of the surplus promised and sustain the match. The attractiveness of unemployment increases in the probability of meeting a new employer $(\mathbb{M}_{u,t})$ and the expected surplus associated with this meeting. The wage falls as the potential gains from a job-to-job transition increase. In this manner, the employer extracts some of the surplus gain from potential job-to-job moves. An increase in potential job-to-job transition gains arises due to either a higher meeting rate on-the-job $(\mathbb{M}_{e,t})$, better job prospects in terms of potential matches $(1 - \rho)$, higher future surplus at the current job $(\widehat{S})$, or higher expected future surplus elsewhere $(\bar{S})$. Deteriorating job prospects for the worker due to a fall in $1 - \rho$, $\bar{S}$, or $\widehat{S}$ increase the wage.

(13) does not yield unambiguous predictions for wage changes at the individual or aggregate level if $\mathcal{F}$ changes. Consider an economy with only two occupations with skill requirements $\hat{\mathbf{y}}$ and $\tilde{\mathbf{y}}$, respectively, and a set of workers whose skills are such that surplus is highest with their current employer. Then, a fall in the probability mass on $\hat{\mathbf{y}}$ decreases $1 - \rho$ and $\bar{S}$ and thus increases the wage for a worker at employer $\hat{\mathbf{y}}$. However, the gains from an employment to unemployment transition fall, thus decreasing the wage for this worker and offsetting the increase just described. In this manner, the contraction of an occupation puts upward pressure on wages, but wages in the occupation may rise or fall. The wage effect for an individual depends on which change in the continuation value dominates. Naturally, the effect on average wages within an occupation depend on the distribution of the individuals within the occupation and how their individual

changes aggregate.

2.4.1 Equilibrium Concept

We can now consider an equilibrium for this model. I focus on an equilibrium concept where the economy transitions from one steady state in 1979 to another in 2010. This equilibrium path is the outcome of decentralized, optimal individual behavior over time given beliefs about objects that change over time and taking others behavior as given. A transition path equilibrium allows changes in skill requirements and productivity over time, which generate changes in the equilibrium wage distribution and occupational structure. Skill requirements or demands, $\mathcal{F}_t(\mathbf{y})$, evolve over time to produce job polarization in this model.[32] Productivity evolves (f_t) over time and contributes to changes in wage outcomes and changes in the occupational structure through sorting (also referred to as selection effects).

In this model, workers and employers must form beliefs over how these skill demands will evolve in order to make decisions about what matches to form and determine wages. The two most straightforward albeit extreme cases are perfect foresight and no anticipation. Under perfect foresight, all agents know the entire path $\{\mathcal{F}_t(\mathbf{y})\}_{t=0}^T$ and $\{f_t(\mathbf{x},\mathbf{y})\}_{t=0}^T$ (i.e. z_t) following an unanticipated change at time 0. Under no anticipation, changes in z_t surprise all agents each period and z_t remains their best guess of z_{t+1}. Comparing these cases provides insight on the importance of expectations over the future demand for an occupation in job selection and wage determination.

In Appendix A.5, I define the general rational expectations equilibrium and explain the difficulties in solving for it outside of a steady state.[33] I then make the case for this more restrictive but more easily solved partial equilibrium, which I use to take the model to the data. Here, I provide the definition. A partial equilibrium must consist of the solutions to (5), (6), and (9) which characterize equilibrium wages (13) given that free entry assumption drives equilibrium V_t (7) to zero.

Definition 2.1 (Partial Equilibrium Path).
Given $\{z_t\}_{t=0}^T$, the tuple $\{U_t(\mathbf{x}), W_t(\mathbf{x},\mathbf{y},\sigma), P_t(\mathbf{x},\mathbf{y},\sigma), V_t, w_t(\mathbf{x},\mathbf{y},\sigma)\}$ form a partial equilibrium path from time 0 to time T if the following hold.

1. (5), (6), and (9) solve $U_t(\mathbf{x})$, $W_t(\mathbf{x},\mathbf{y},\sigma)$, and $P_t(\mathbf{x},\mathbf{y},\sigma)$, respectively
2. $w_t(\mathbf{x},\mathbf{y},\sigma)$ satisfies (13) for all employed workers
3. $V_t = 0$ at every period t by (8) [Free Entry]
4. Agents hold beliefs over the path of $\{z_t\}_{t=0}^T$, i.e. $\{\mathcal{F}_t(\mathbf{y})\}_{t=0}^T$ and $\{f_t(\mathbf{x},\mathbf{y})\}_{t=0}^T$

Solving this equilibrium amounts to backwards solving (12) from (11) at time T back to time 0 when the unanticipated changes to z_t hit. If the agents' beliefs coincide with the actual paths of $\mathcal{F}_t$ and f_t, then it can be a considered a rational partial equilibrium path.

3 Data

Use of the task framework became popular with Autor, Levy, and Murnane well over a decade ago. Naturally, the datasets used to analyze the task content of occupations in the US are well-known, well-documented, and

[32]But not just job polarization. $\mathcal{F}_t(\mathbf{y})$ also affects the wage distribution.
[33]Exogenous meeting rates make the equilibrium partial. Appendix A.5 endogenizes the meeting rates to show the general equilibrium.

widely used now. These datasets include the Current Population Survey (CPS) and National Longitudinal Survey of Youth (NLSY) for workforce data over time and the Dictionary of Occupational Titles (DOT) and the Occupational Information Network (O*NET) for task content and complexity information. Estimating the model requires time-varying information on hourly wages, employment shares and wages across occupations, the equilibrium distribution of $\mathbf{y}$, and the distribution of initial worker skill endowments. The aforementioned datasets provide a means to obtain this information.

3.1 Wages and Employment Shares

Changes in the wage distribution and employment shares provide variation to estimate productivity parameters and the distribution of skill requirements. I use the CPS to measure changes across the wage distribution and employment shares from 1979 to 2010. I draw on the Outgoing Rotation Group (ORG) of the CPS to do so. The CPS ORG consists of roughly a quarter of the monthly CPS administered by the US Census Bureau. The Bureau interviews households for 4 months, rotates them out of the survey for 8 months, and rotates them back into the survey for a final 4 months. The ORG consists of individuals interviewed in the last month of each rotation and provides point-in-time measurements of wages for most workers. The March CPS Annual Social and Economic Supplement (ASEC) provides household income and demographic data used extensively to study income inequality.[34] However, the March supplement does not provide point-in-time measures of hourly wages in contrast to the CPS ORG questions. This point-in-time measurement makes the quality of wage data in the ORG considerably higher than the ASEC (Lemieux, 2006). From the CPS ORG, I pool monthly observations to construct an annual dataset of hourly wages, occupations, and demographic information. I provide detailed information on dataset construction, occupational harmonization, sample restrictions, and summary statistics in Appendix B.1.

I estimate the model at the level of hourly wages for several reasons. First, the model does not have an intensive margin with respect to labor supply (i.e. hours worked). Hourly wages better reflects changes in productivity and skill requirements solely due to changes on the extensive margin. Second, most workers (approximately 60%) in the economy receive hourly pay rates. The number of workers receiving hourly pay rates has also remained stable around 60% since 1979.[35] Examining wages and not total compensation raises the concern that perhaps changes in non-wage benefits rather than productivity or skill requirements explain changes across the wage distribution. The share of non-wage compensation has arguably increased as shown in Sherk (2013), however the vast majority (approximately 80%) of total compensation still consists of wages according to macro-level data. Available evidence suggests wages well reflect changes in the economic returns to a job even in the presence of non-wage benefits (Katz and Autor, 1999). The lack of extensive individual-level data on compensation composition over time makes it difficult to say whether increased non-wage benefits account for wage polarization and expansion. Katz and Autor (1999) argue from available evidence at the time that non-wage benefits actually tend to reinforce rather than offset changes in the wage distribution we observe over the 80s and 90s.[36] Thus, I judge hourly wages to be a fair metric on which to

[34] See Heathcote, Perri, and Violante (2010) for example.

[35] See Table 10 of Bureau of Labor Statistics (2016).

[36] Thus, implementing this model with wage data is not necessarily misleading despite missing out on developments that affect non-wage compensation like healthcare costs. Wages still contain information on productivity and skill requirement developments, which I make inference on using indirect inference as opposed to an exact identification strategy. One could argue that employers and workers only care about total compensation and not wages. Therefore, wages only carry partial information regarding pay and job selection. This argument rests on the assumption that the employer can fully adjust the composition of total compensation. If benefits come in standardized packages for example, then the employers will not be neutral to the wage-benefit composition at the individual level. Recent evidence from Eriksson and Kristensen (2014) suggests employers as well as employees face a nontrivial trade off in determining wages and non-wage benefits. The presence of such a

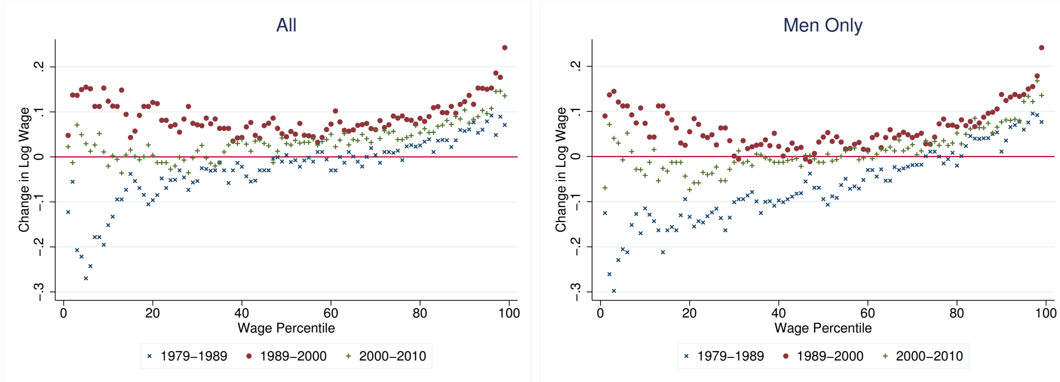


Figure 1: Wage Percentile Changes (1979 to 2010)

examine job selection and pay decisions.[37]

Figure 1 shows changes at wage percentiles across the three decades I consider. The left panel shows all workers and the right panel shows only men. The patterns for men and women appear to differ slightly with more wage growth for women, however an overall picture emerges. Wages expanded across the top wage distribution over all three decades. Wages compressed at the bottom of the wage distribution in the 1990s, which we commonly refer to as wage polarization. In the 2000s, some wage compression appears the very bottom of the distribution, but overall wages did not expand or compress in the bottom half. This figure confirms much of previous findings with respect to wage changes over these decades (Mishel, Schmitt, and Shierholz, 2013).

I present occupational employment share and average wage changes using Acemoglu and Autor's (2011) broad grouping of occupations in Figure 2.[38] They group occupations into low-paid, medium-paid, and high-paid categories.[39] The low-paid category consists only of low-paid service occupations. The medium-paid category consists of sales, clerical, administrative support, production, craft, repair, and operative occupations. The high-paid category consists of managerial, professional and technical occupations. The top panel of Figure 2 confirms the existing evidence regarding job polarization. We observe a relative decline in middle-paid occupations throughout the 1980s and 1990s (Mishel, Schmitt, and Shierholz, 2013; Lefter and Sand, 2011). We also see disproportionate growth in low-skilled service occupations in the latest decade (Autor and Dorn, 2013). The right panel of Figure 2 shows only men. They exhibits similar employment share patterns.

Average wages within occupations diverge from employment share patterns. The gap in average wages expands between occupations in the 1980s as wages overall spread out in the 1980s (Figure 1). Average wages polarize like wages overall in the 1990s as they rise less in the middle-paid occupational group than the low and high-paid groups. However, women appear to drive this pattern as men show less wage growth in the low-skilled occupational group. Average wage differences appear to expand in the 2000s, although not nearly as strongly as in the 1980s. Looking only at men, we observe overall wage polarization even though

tradeoff increases the importance of wages with respect to what information they carry.

[37] The availability of individual wage data compared to total compensation facilitates its widespread use and use here. However, the debate as to whether wages are a sufficient metric to track the evolution of returns to employment remains a contested and important area of research.

[38] Several papers in the job polarization literature present these figures in terms of occupational skill ranks using average wages in a reference year to rank occupations (Acemoglu and Autor, 2011; Autor and Dorn, 2013; Mishel, Schmitt, and Shierholz, 2013). I replicate and discuss these figures in Appendix B.6.

[39] In general, pay reflects skill level, so the literature often uses low-paid and low-skilled interchangeably.

average occupational wages for men do not on strongly polarize between groups. This observation suggests men in the middle-paid group moved down the wage distribution by the start of the 1990s, and their partial wage recovery may account for some of the wage polarization we observed.

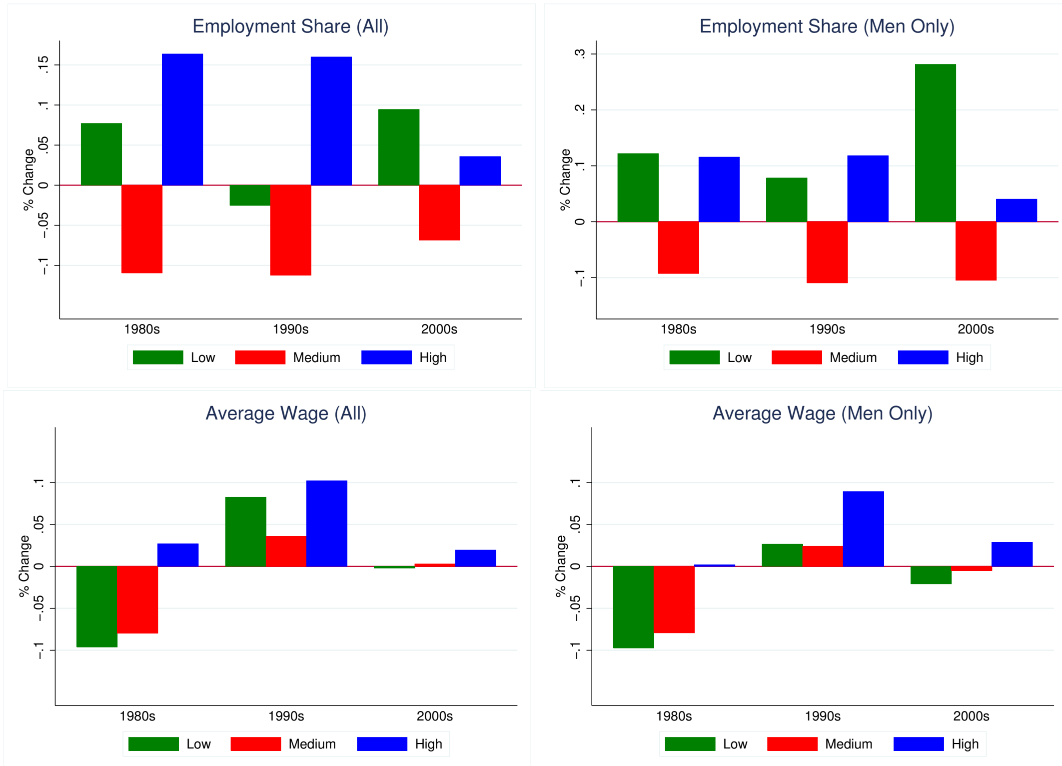


Figure 2: Employment Share and Average Wage Changes (1979 to 2010)

These occupational categories consider in Figure 2 do not yet map into the model primitive regarding skill demands ($\mathbf{y}$). I make use of task content/complexity data to map the occupational data to skill requirements.

3.2 Skill Requirements

The DOT (1977) provides measures of task complexity along cognitive and manual dimensions of skills at the occupation level. Importantly, the DOT features information gathered from direct observation of the tasks performed in an occupation and thus measures task complexity independent of worker skills.[40] The US Department of Labor infrequently updated the DOT before replacing it with the Occupational Information Network (O*NET) in the late 2000s. Even so, the DOT remains relevant to most of the period under consideration. I manually update DOT task information using O*NET due to the emergence of new occupations in 2003 where the DOT does not provide information. More recent work like Lise and Postel-Vinay (2016) use O*NET instead of the DOT. I compare the DOT and O*NET and argue the case for using the DOT here in Appendix B.4. Both the DOT and O*NET come with the severe drawback that they only capture task changes between but not within occupations.[41] More recent attempts at analyzing task content within occupations include the Occupational Requirements Survey (ORS) and Autor and Handel (2013).

[40]It would not be credible to use measures conditional on the distribution of worker skills to construct the equilibrium distribution of $\mathbf{y}$, because this (unobserved) distribution of worker skills changes over time as well.

[41]I provide more detailed information about the DOT and O*NET datasets and their drawbacks in Appendix B.2 and B.4.

The ORS only began releasing data in late 2016 (Bureau of Labor Statistics, 2017). Autor and Handel collect representative survey data that allows them to capture differences in task content within occupations at a point in time. Given limitations to data availability, the DOT provides an acceptable and widely used means to obtain task measures.

I use the Dictionary of Occupational Titles combined with the CPS ORG to estimate equilibrium skill requirements. Many papers like Autor, Levy, and Murnane (2003) and Yamaguchi (2012) use this merging method and data. They also provide detailed descriptions and discussions of the DOT. I relegate those descriptions and discussions to Appendix B.2 and B.4 and focus on the main procedures here. I merge DOT measures into the CPS ORG using Dorn's harmonized occupational coding system (Dorn, 2009). This combined CPS-DOT dataset contains DOT task measures on the occupational level and individual weights to construct skill scores. From here, I construct cognitive and manual skill scores using principle components analysis à la Yamaguchi (2012). I use general learning ability, verbal ability, and numerical ability to estimate cognitive skill requirements. I use an array of other aptitudes to measure manual skill requirements, including physical strength, motor coordination, finger dexterity, and manual dexterity. Appendix Table 13 contains details on all these additional measures. I take the first principle component in each case and linearly rescale it to the interval $[0, 1]$.[42] Some papers convert these task measures to percentiles, however this transformation makes all occupations equidistant. I preserve the distance in skill requirements between occupations, because this distance governs differences in output and consequently differences in wages between occupations.

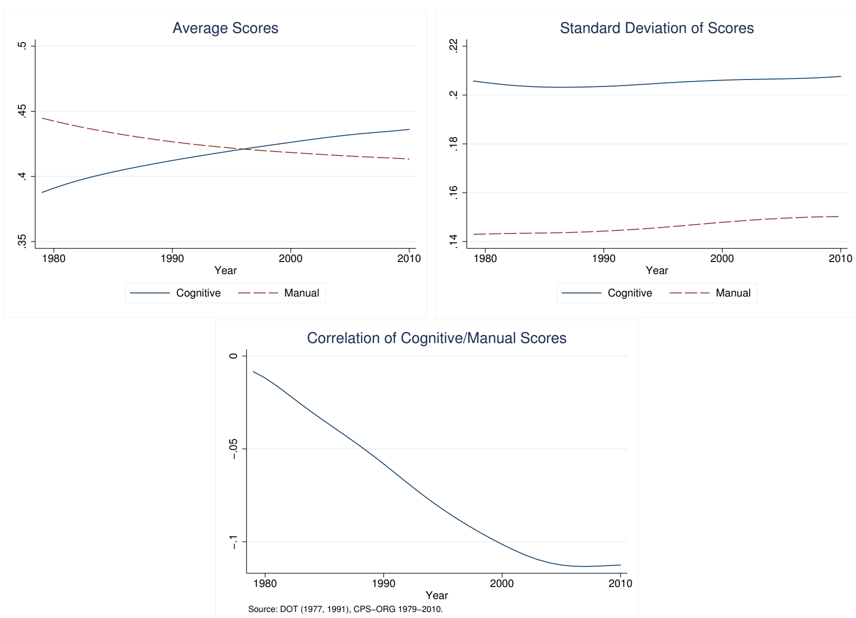


Figure 3: DOT Equilibrium Skill Requirement Moments (1979 to 2010)

Figure 3 shows the resulting moments.[43] We observe a rise (fall) in the mean level of cognitive (manual) task complexity and increased dispersion in both dimensions of task complexity. In addition, the (negative)

[42] I could also use two exclusion restrictions and the first two principle components to identify cognitive and manual skills. I implement this alternative approach, rotating the first two principle component scores based on the restriction that general learning ability and motor coordination reflect only cognitive and manual skill, respectively. This alternative approach yields cognitive skills with a correlation each year of at least 0.99 for cognitive skills and 0.96 for manual skills. Thus, the approach makes little difference with respect to the final skill scores. I use the apporach of running two separate factor analyses for ease of interpretation.

[43] I smooth the time series of the moments to reduce sampling noise using Lowess with the optimal bandwidth.

correlation between cognitive and manual task complexity falls until the last decade. I also show mean skill requirements for the major occupational and industry groups in Tables 1 and 2 from 1979 to 2010. The scores appear intuitive, and the following hold on average. High-skilled (management, professional, technical) occupations require the most complex cognitive tasks and thereby the most cognitive skills. Low-skilled service occupations require the least cognitively complex task and thus the least cognitive skills. The middle-skilled occupations (clerical to products and crafts) require varying amounts of cognitive task complexity but relatively higher manual task complexity compared to low and high-skilled occupations. This feature suggests that technological change inducing middle-skilled occupations to contract works through eliminating manually complex tasks. Industries requiring the highest cognitive task complexity include financial services, professional and business services, and educational and health services. Industries requiring the highest manual task complexity include manufacturing, construction, and mining. The service industry requires the lowest levels of cognitive task complexity.

Table 1: Mean Skill Requirements by Major Occupational Group

	y_C	y_M
Management, Professional, Technical	0.613	0.384
Clerical and Retail Sales	0.427	0.417
Construction, Mechanics, Mining, Transport	0.264	0.525
Machine Operators, Assembling, Inspection	0.370	0.543
Products and Crafts	0.161	0.480
Service	0.193	0.385

Table 2: Mean Skill Requirements by Major Industry Group

	y_C	y_M
Agriculture, Forestry, Fishing, and Hunting	0.312	0.446
Mining	0.382	0.462
Construction	0.363	0.517
Manufacturing	0.344	0.467
Wholesale and Retail Trade	0.374	0.398
Transportation and Utilities	0.337	0.440
Information	0.479	0.418
Financial Services	0.548	0.350
Professional and Business Services	0.510	0.402
Educational and Health Services	0.471	0.428
Leisure and Hospitality	0.342	0.389
Other services	0.290	0.465

The mean skill requirements for these major occupational groups also suggests a simple mapping from skill requirements $\mathbf{y}$ to low, middle, and high-skilled occupational groups. Figure 4 plots average $\mathbf{y}$ for all occupational titles 1979 to 2010 with red lines at 0.4 on the x-axis and 0.60 and 0.45 on the y-axis. This figure and Table 1 suggest cutoffs in the level of cognitive and manual skill provide a fair mapping from skill requirements to occupational categories. I consider the breakdown where jobs with $y_M < 0.4$ and $y_C < 0.45$ make up low-skilled occupations, jobs with $y_M \geq 0.4$ and $y_C < 0.6$ make up middle-skilled and the rest are high-skilled. This breakdown, weighting occupations by their employment share, captures approximately 70%

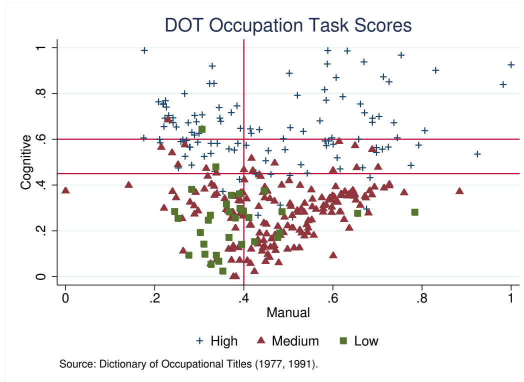


Figure 4: Average Skill Requirements by Occupation

of the employment for each category (low, middle, high) using the Acemoglu and Autor (2011) occupational groups. Coincidently, this particular breakdown nearly reaches the cutoffs that best match the Acemoglu and Autor categories. Occupational titles and task content define the Acemoglu and Autor categories whereas only task complexity defines these occupational categories. Yet, the two groupings overlap significantly, which suggests task complexity captures a lot of information about occupations. Many occupational mappings based on skill requirements are possible, however none match the simplicity and intuitive appeal of this one. With this mapping, the relative contraction of manually complex tasks corresponds to the relative contraction of middle-skilled occupations. The relative expansion of cognitively complex tasks corresponds to the relative expansion of high-skilled occupations.

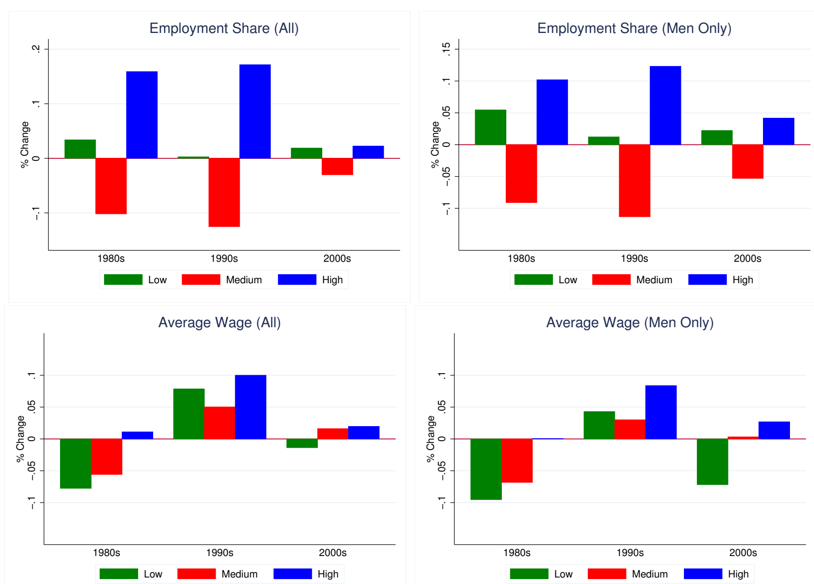


Figure 5: Occupational Changes based on Skill Requirement Definition

I replicate Figure 2 using the occupation mapping described. Unsurprisingly, Figure 5 and Figure 2 exhibit similar overall patterns. These groupings overlap substantially, so they should look similar. However, they differ in levels. The categorization based on skill requirements results in more low-skilled occupations[44], thus

[44]21.8% vs. 11.8% in 1979, 23.3% vs. 16.3% in 2010. Examples of occupations recategorized into the low-skilled category by skill requirements include hotel clerks and parking lot attendants. Most occupations recategorized from middle-skilled to

lowering the increase in the share of low-skilled occupations in the 2000s. It also results in less high-skilled occupations[45], thus raising the increase level for high-skilled occupations. In addition, male occupational wages polarize more under the skill requirement categorization unlike in Figure 5.

3.3 Worker Skills

The National Longitudinal Survey of Youth (1979) provides nationally representative information to construct the distribution of entering worker skill endowments, $\mathcal{V}_t(\mathbf{x})$. It also provides observations on the joint distribution of worker skill endowments and skill requirements, labor market transitions, and wages. I use the NLSY to construct $\mathcal{V}_0(\mathbf{x})$ as well as estimate some micro-level moments requiring panel data. Much of my treatment and construction of the NLSY parallels Lise and Postel-Vinay (2016) and Boehm (2017). I elaborate on this construction and its issues in more detail in the Appendix B.5.1 and describe the main process here.

I construct $\mathbf{x}(0)$ analogously to skill requirements. I use the Armed Services Vocational Aptitude Battery (ASVAB) test in the NLSY79, which provides pre-labor market entry scores for mathematics knowledge, arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, general science, coding speed, auto and shop information, mechanical comprehension, and electronics information (Bureau of Labor Statistics, U.S. Department of Labor, 2014a). I extract the first two principle components of all the ASVAB scores, and impose two exclusion restrictions to identify cognitive and manual skill scores. I restrict mathematical knowledge to contain information only on cognitive skills, and auto and shop information to contain information only on manual skills. I linearly rescale these scores into the interval $[0, 1]$ to form estimates $(\widehat{x}_C(0), \widehat{x}_M(0))$. For worker skills, I rotate the first two principle component scores instead of separating the measures into categories, because the ASVAB measures do not categorize as easily as the DOT measures. Tests about mechanical comprehension and electronics likely convey information about cognitive and manual skills as both tests require some knowledge of general science and reading comprehension.

Figure 6 shows the constructed $\widehat{\mathcal{V}}_0(\mathbf{x})$ for the NLSY1979 cognitive and manual skills.[46] It also shows $\widehat{\mathcal{V}}_0(\mathbf{x})$ conditional on gender and educational attainment groups. The distribution of initial cognitive skills across education groups appear intuitive. Higher education groups exhibit more cognitive skills. They also tend to exhibit more manual skills although to a much lesser difference than cognitive skills. Initial cognitive skills across gender do not notably differ, while initial higher manual skills exhibit a strong skewness towards males. I reweigh $\widehat{\mathcal{V}}_0(\mathbf{x})$ using the observed educational attainment and female share of the labor force to obtain $\widehat{\mathcal{V}}_t(\mathbf{x})$ over time. This approach remains sensible only if the distribution of cognitive and manual skills remains similar within education-gender cells of cohorts. In Appendix B.5.1, I use the NLSY97 to validate this restriction, which shows little difference between $\widehat{\mathcal{V}}_{1979}(\mathbf{x})$ and $\widehat{\mathcal{V}}_{1997}(\mathbf{x})$ within education-gender group for cognitive skill but more difference for manual skills. Finally, I allow a transformation of $\widehat{\mathbf{x}}(0)$ into $\mathbf{x}(0)$ in the estimation to align it with $\mathbf{y}$, because $\widehat{\mathbf{x}}(0)$ need not necessarily align with the DOT $\mathbf{y}$.

low-skilled occupations are lower level clerical or manufacturing occupations.

[45] 29.0% vs. 25.3% in 1979, 42.0% vs. 36.2% in 2010. Nearly all recategorized occupations are from high-skilled to middle-skilled occupations. They mainly consists of health and human service related occupations like nursing, occupational therapists, physical therapists, and clinical technicians.

[46] Workers initial ages vary based on educational attainment level.

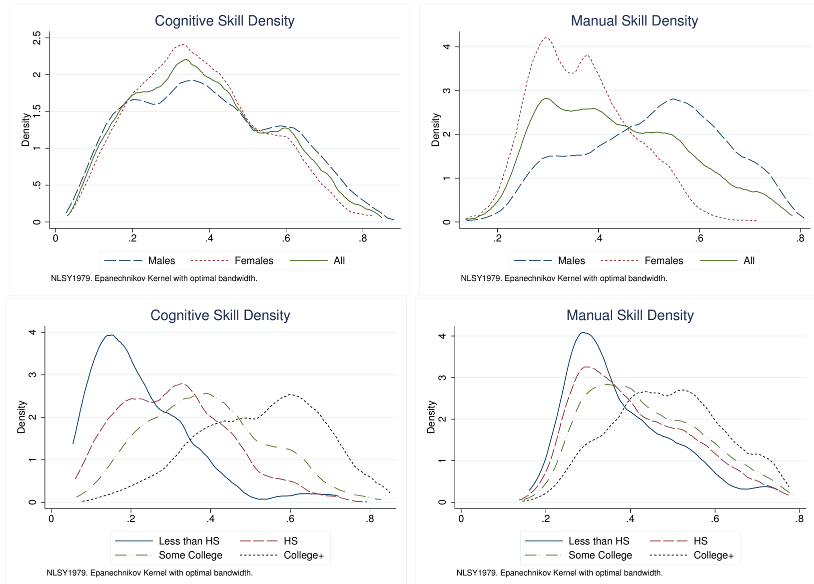


Figure 6: Marginal Distributions of Initial Worker Skills

4 Estimation

I estimate it via indirect inference after parameterizing it.[47] The model requires parameterization for the following objects: $\mathcal{F}_t(\mathbf{y})$, $f_t(\mathbf{x}, \mathbf{y})$, $c(\mathbf{x}, \mathbf{y})$, $b(\mathbf{x})$, and $h(\mathbf{x}, \mathbf{y})$. It also requires estimating or calibrating β_a, δ, ω, λ, $\mathbb{M}_e$, and $\mathbb{M}_u$. Indirect inference requires three steps. First, I set parameter values. Second, I solve the model. Given parameter values, we know $f_t(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y}) - b(\mathbf{x})$. This function corresponds to the surplus function in the final period before mandatory retirement. I solve the model backwards from this terminal period in the worker's life. Third, I simulate the model to produce the targeted moments. Estimation iterates over this process until the model suitably reproduces the targeted moments.[48]

4.1 Parameterization

The parameterization I employ relies heavily on the one developed in Lise and Postel-Vinay (2016), because it yields a good fit to many aspects of the data. The production function consists of linear terms in skill requirements and quadratic terms to capture complementarities and under-qualification in a skill dimension. Skill requirements reflect the employers production technology and thereby directly affect output. In turn, output loss in production drives positive sorting across task dimensions as it prevents matches with severely under-qualified workers in some task dimension. The degree of output loss increases convexly with the distance between skill and skill requirements. The wage function (13) reflects the curvature in the production function. Matching changes in the shape of the wage distribution may require a change in the convexity in the production function. To this end, I introduce within-task complementarity terms $x_C y_C$ and $x_M y_M$ to form the production function in (14).[49] This production function exhibits absolute advantage in productivity

[47]Lise and Postel-Vinay (2016) provide a formal identification argument for their model based on specific functional forms that yield a closed form solution for the surplus function (see their Appendix A.6). The exact argument remains too stylized to apply directly here, however the spirit of their identification argument holds relevance for which features of the data to target. Non-parametric identification of static search models is an emerging area of research, but it has yet to be extensively addressed for dynamic models.

[48]I estimate based on the equally weighted minimum distance loss function.

[49]Lindenlaub (2017) permits such within-task complementarities and shuts down between-task complementarities.

as excess skills do not hurt output. General skills amplify output, which magnifies differences within cohorts over time.[50]

$$f(\mathbf{x},\mathbf{y}) = x_G \cdot \Big[\alpha_{0,t} + \alpha_{C,t}y_C + \alpha_{M,t}y_M + \alpha_{CC,t}x_Cy_C + \alpha_{MM,t}x_My_M \\ - \kappa_C \min\{x_C - y_C, 0\}^2 - \kappa_M \min\{x_M - y_M, 0\}^2\Big] \tag{14}$$

Empirically, over-qualified workers experience wage loses compared to workers with similar skill levels positioned in jobs where their skills are required. They also receive higher wages compared to workers with just-qualified skills (i.e. $\mathbf{x} = \mathbf{y}$), doing the same job (Slonimczyk, 2013). The disutility of labor serves to permit these empirical observations. Workers only experience labor disutility in dimensions of over-qualification as shown in (15).[51] General skills amplify the effect of labor disutility. I specify the flow utility of an unemployed worker with the same general structure as the production function shown in (16). However, it does not depend on specific skills.[52]

$$c(\mathbf{x},\mathbf{y}) = x_G \cdot \Big[\nu_C \cdot \max\{x_C - y_C, 0\}^2 + \nu_M \cdot \max\{x_M - y_M, 0\}^2\Big] \tag{15}$$

$$b(\mathbf{x}) = x_G \cdot b_0 \tag{16}$$

The worker's specific skills accumulate or depreciate linearly in task dimensions shown in (17). This learning-by-doing specification varies the skill acquisition or loss according to the distance between the worker's current skill and skill requirements. In this manner, learning-by-doing is heterogeneous across workers.

$$h(\mathbf{x},\mathbf{y}) = \underbrace{\mathbf{x}}_{\text{skill today}} + \underbrace{\Gamma_H \cdot \max\{\mathbf{y} - \mathbf{x}, \mathbf{0}\}}_{\text{skill gain}} + \underbrace{\Gamma_D \cdot \max\{\mathbf{x} - \mathbf{y}, \mathbf{0}\}}_{\text{skill depreciation}} \tag{17}$$

$$\Gamma_H = \begin{pmatrix} \gamma^h_{CC} & \gamma^h_{CM} \\ \gamma^h_{MC} & \gamma^h_{MM} \end{pmatrix}, \quad \Gamma_D = \begin{pmatrix} \gamma^d_{CC} & \gamma^d_{CM} \\ \gamma^d_{MC} & \gamma^d_{MM} \end{pmatrix}$$

I specify general skills (18) as a function of age and an individual component (ε) weighted by initial cognitive skills. This component exists to capture wage dispersion in talent among workers in the most cognitive-intensive jobs. The quadratic age term serves to capture the wage-age/experience trend.

$$x_G(t) = \gamma_0 + \gamma_1 age(t) + \gamma_2 age^2(t) + x_C(0) \cdot \varepsilon \tag{18}$$

Finally, I parameterize $\mathcal{F}_t(\mathbf{y})$ using five time-varying parameters (r_t, $a_{C,t}$, $b_{C,t}$, $a_{M,t}$, $b_{M,t}$) to characterize the sampling distribution of skill requirements. The time-varying nature of the parameters makes parsimony crucial to estimate the model. To this end, I use a Clayton copula to characterize the joint distribution of y_C and y_M. It consists of one parameter (r) controlling the correlation between y_C and y_M.[53] I use Kumaraswamy marginals for y_C and y_M. This marginal provides a closed form cumulative distribution function over the support $[0, 1]$, making it more tractable.[54] Each marginal consists of two parameters

[50] Empirically, wages exhibit increasing dispersion at higher ages.

[51] Over-qualification does not cause output loses, but it does lower the total surplus of a match thus indirectly lowering wages. Thus, over-qualified hold an absolute advantage but still may be undesirable. Over-qualification also increases wages in some jobs as the wage equation (13) shows. The disutility of labor due to over-qualification enters wages positively as the employer compensates the worker for said disutility. In this manner, over-qualified workers can receive higher wages compared to workers with just-qualified skills in the same job but lower wages relative to others in their skill level in more skilled jobs.

[52] I impose this restriction to reduce the number of parameters to estimate.

[53] This copula falls into the class of Archimedean copulas, and its closed form conditional distribution function simplifies the sampling process.

[54] Kumaraswamy approximates a Beta distribution and can be shown to map into a generalized Beta distribution (Jones,

governing the shape of the marginal. The first shape parameter pushes mean and variance in opposite directions, while the second pushes them in the same direction. This feature makes the model able to match similar trends with respect to the mean and dispersion of cognitive task complexity but opposing trends for marginal task complexity.

The five parameters for $\mathcal{F}_t(\mathbf{y})$ and four parameters in $f_t(\mathbf{x}, \mathbf{y})$ vary over time, and I specify a process for how they evolve. One obvious approach estimates these parameters at each point in time. I estimate the model at the monthly level over 32 years, rendering this approach intractable. Instead, I allow them to evolve over time using linear time trends with structural breaks to capture different trends across decades. Time trends with structural breaks provide a compromise of flexibility and parsimony. I set structural breaks to occur near the start and end points of each decade.[55]

4.2 Solution and Simulation Protocol

Given parameters, I solve for the match surplus function and simulate a model labor market to produce the targeted simulated moments. I simulate the labor market monthly for approximately 50,000 workers from January 1979 to December 2010. The simulation consists of a burn-in period, a transition period, and a terminal period. I index the transition period January 1979 as time period 1 ($t = 1$) and December 2010 as time period 384 ($t = 384$). I index the burn-in period as $t = 0$ and the period after December 2010 as $T = 385$. $\mathcal{F}(\mathbf{y})$ and $f(\mathbf{x}, \mathbf{y})$ do not vary over time during the burn-in period or after the transition.[56]

To solve the model, I first solve the model before (burn-in) and after the transition. $S(\mathbf{x}, \mathbf{y})$ equals the static portion of the match surplus, $f(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y}) - b(\mathbf{x})$, for workers whose age next period is 65. Given $S(\mathbf{x}, \mathbf{y})$ at this terminal age, I exploit the recursive structure of the surplus function and solve backwards over age to obtain $S_0(\mathbf{x}, \mathbf{y})$ and $S_T(\mathbf{x}, \mathbf{y})$. Next, I use the recursive structure of the surplus again to solve $S_t(\mathbf{x}, \mathbf{y})$ backwards over time from $t = 384$ to $t = 1$. The perfect foresight solution uses the time-varying parameters at their respective times. In contrast, the no foresight solution does not incorporate information from the future. In the case of no foresight, the agents assume no parameters vary over time, i.e. $S_{t+1}(\mathbf{x}, \mathbf{y}) = S_t(\mathbf{x}, \mathbf{y})$. Thus, obtaining $S_t(\mathbf{x}, \mathbf{y})$ requires solving the model backwards over age at every point in time $t = 0, 1, ..., 384, 385$.

Given the surplus function, the simulation protocol produces a cross section of worker skills ($\mathbf{x}_{it}$), skill requirements ($\mathbf{y}_{it}$), surplus shares (σ_{it}), and labor market transitions. From here, I construct wages based on (13), employment shares based on the mapping in Figure 4, and labor market transition rates. I add a zero-mean, log-normal measurement error with standard deviation υ to simulated wages, because the data exhibits measurement error. The simulation protocol starts with a burn-in period, holding all parameters fixed. To initialize the burn-in period, all workers start out employed at a random $\mathbf{y}_{i0}$ and draw skills $\mathbf{x}_{i0}$ from $\mathcal{V}_0(\mathbf{x})$.[57] Matches immediately terminate where the surplus is negative. The simulation then runs through the burn-in period to the terminal period ($t = 385$).

One period (t) of the simulation goes as follows. A worker starts the period with skills $\mathbf{x}_{it}$ aged a_{it}. The worker exits the labor force and the match terminates with probability ξ_a. An employed worker's match with $\mathbf{y}_{it}$ terminates with probability δ. An employed worker in a surviving match meets another employer with probability $\mathbb{M}_e$. An employed worker who meets a new employer then draws $\mathbf{y}'$ from $\mathcal{F}_t(\mathbf{y})$ and moves

2009).

[55] I also explored including the dates of the breaks in the optimization routine. However, they did not change much from around the start and ends of decades naturally, because the timing of targeted moments is decadal.

[56] Obviously, this approach misses out on any forward looking effects from the 2010s, which may affect decisions and wages in the 2000s. However, it provides a clean way to estimate the transition path. In the case of no foresight, this issue is irrelevant.

[57] I burn-in this labor market for 1000 periods, which provides enough time for the initial cohort of workers randomly assigned to jobs to exit the labor market. I draw workers initial ages from the 1979 cross-sectional age distribution in the CPS ORG.

to the new employer if $S_{t+1}(\mathbf{x}', \mathbf{y}') > S_{t+1}(\mathbf{x}', \mathbf{y}_{it})$ where $\mathbf{x}' = h(\mathbf{x}_{it}, \mathbf{y}_{it})$. A worker who accepts starts at the new employer next period and the surplus share (σ_{it}) updates according to (4). An employed worker who fails to meeting a new employer draws $\mathbf{y}'$ from $\mathcal{F}_t(\mathbf{y})$ with probability ω. Matches with new skill requirements tomorrow terminate if $S_{t+1}(\mathbf{x}', \mathbf{y}') < 0$. In the case of separation, a worker starts the next period unemployed. An unemployed worker meets an employer with probability $\mathbb{M}_u$. An unemployed worker who meets an employer then draws $\mathbf{y}'$ from $\mathcal{F}_t(\mathbf{y})$ and moves to the employer if $S_{t+1}(\mathbf{x}', \mathbf{y}') \geq 0$ where $\mathbf{x}' = h(\mathbf{x}_{it}, \mathbf{0})$. A unemployed worker who accepts starts at the new employer next period and the surplus share (σ_{it}) equals λ. Otherwise, that worker starts next period unemployed. A worker out of the labor force enters at the start of the period.[58] This worker draws new skills from $\mathcal{V}_t(\mathbf{x})$ and searches as an unemployed worker.[59]

4.3 Target Moments

The model consists of two sets of parameters − time varying and time invariant. Time invariant parameters include κ_C, κ_M, ν_C, ν_M, b_0, $(\gamma_0, \gamma_1, \gamma_2)$, Γ_h, Γ_d, $\tilde{\beta}$, δ, ω, λ, ξ_a, (θ_0, θ_1), (ζ_C, ζ_M), arrival rates $\mathbb{M}_e$ and $\mathbb{M}_u$, and measurement error variance υ^2.[60] The ζ parameters map the initial skills estimates $(\widehat{x}_C(0), \widehat{x}_M(0))$ into $(x_C(0), x_M(0))$ via $x(0) = \widehat{x}(0)^{\zeta}$ to better align $\mathbf{x}$ and $\mathbf{y}$. The θ parameters are the scale and shape parameters for Pareto-distributed individual heterogeneity ε. Time varying parameters include the five parameters of $\mathcal{F}_t(\mathbf{y})$, $\alpha_{0,t}$, $\alpha_{C,t}$, $\alpha_{M,t}$, $\alpha_{CC,t}$, and $\alpha_{MM,t}$. I calibrate some parameters externally and estimate the others using variation in the data.[61]

4.3.1 External Calibration

I fix a small number of parameters and show these externally calibrated parameters in Table 3. I set the monthly discount factor $\tilde{\beta}$ as commonly done in the literature. Its value roughly corresponds to a 10% steady state discount rate per annum. I add (zero-mean, log-normal) measurement error to wages as occurs in the CPS ORG wage data. Lemieux (2006) measures the variance of measurement error in wages in the CPS ORG. I set υ^2 to around the level estimated there. I set the involuntary separation probability δ to its counterpart in the data. IPUMS-CPS identifies voluntary and involuntary unemployment, and I apply Shimer (2012) to construct monthly worker flows. I set δ to match the monthly involuntary flow from employment to unemployment. Similarly, I estimate ω to match the involuntary flow into unemployment from employment. Finally, I calibrate entry (μ_a) and exit (ξ_a) probabilities as a function of age to match age-based transition rates in and out of the labor force.

[58] Lise and Postel-Vinay (2016) only simulate a cohort of workers to focus on the origins and costs of skill mismatch whereas I simulate a model labor market, allowing new cohorts to enter.

[59] The worker draws an education level and a gender and then draws an age, ε, and cognitive and manual skills the education-gender-group distribution of $\widehat{\mathcal{V}}_0(\mathbf{x})$. Workers initial ages vary based on educational attainment level. No population growth or shrinkage occurs, so new workers enter when old workers exit.

[60] I stick to a partial equilibrium, restricting $\mathbb{M}_{e,t}$ and $\mathbb{M}_{u,t}$ to remain exogenous and time invariant. These parameters vary in the general equilibrium as shown in Appendix A.5. In general equilibrium, these rates vary with the endogenous distribution of worker types. The need for individual agents to forecast and track this endogenous distribution makes the general equilibrium model intractable. We can also interpret this model as an approximation to the general equilibrium outcome where its accuracy depends on the strength of general equilibrium feedback onto the meeting probabilities.

[61] Here, I give the intuition for which variation in the data helps identify the parameters. However, I provide an extensive identification argument for the estimated parameters given a sufficiently rich panel data set in Appendix A.6. This argument further illuminates how the moments targeted in indirect inference help identify the parameters.

Table 3: External Calibration

Parameter	Value	Target
$\tilde{\beta}$	0.992	10% discount rate per annum
δ	0.012	Average Monthly Involuntary Separation Rate
υ^2	0.020	Lemieux (2006)

4.3.2 Estimation Moments

I estimate the remaining parameters jointly. The CPS-DOT provides information for most moments, while the NLSY79 cohort provides information for some of the more micro-level moments. As mentioned, I estimate ω to match the overall separation rate. The δ shock generates the involuntary flows from employment to unemployment. The remaining flows come from voluntary separation following a productivity shock. Thus, ω reproduces the overall average monthly, separation rate conditional on the model's other parameters. Along similar lines, $\mathbb{M}_e$ and $\mathbb{M}_u$ reproduce the monthly job-to-job and unemployment to employment transition rates given the rest of the model parameters. Hence, I target the average monthly job-to-job transition rate, unemployment to employment flow rate and employment to unemployment flow rate.[62]

I target the shape of wage-age profile to pin down values for $(\gamma_0, \gamma_1, \gamma_2)$ in the estimation. I also target the differential between average wages overall and wages out of unemployment to estimate b_0, because b_0 determines wage out of unemployment conditional on the model's other parameters. The wage drop following an unemployment spell contains information on the worker's bargaining power out of unemployment conditional on the model's other parameters like Γ_d. I compute average wage drop following an unemployment spell from the sample NLSY79 panel (Appendix B.5.1) and use it to provide information for the bargaining power λ.

I include moments on the correlation of initial skills and skill requirements at various dates to estimate κ_C, κ_M, ν_C, ν_M, Γ_h, and Γ_d. The correlation of $\mathbf{x}$ and $\mathbf{y}$ in skill dimension measure sorting patterns of worker type $\mathbf{x}$ across jobs with skill requirements $\mathbf{y}$. Parameters κ_C, κ_M, ν_C, and ν_M govern the sorting patterns across worker skill and job skill requirements. For instance, a worker close to zero in the cognitive dimension cannot obtain a job with cognitive task complexity close to one given a high enough κ_C. Similarly, a worker with cognitive skill close to one rejects a job with cognitive requirements close to zero for high enough ν_C. Γ_h also governs sorting patterns. For example, suppose Γ_h is the identity matrix as opposed to the zero matrix. Skills adjust to skill requirements after one period. Severe under-qualification in any skill dimension poses a much lower barrier to obtaining the job in question in this case. Thus, the correlation between initial skills and skill requirements will be low. Intuitively, one does not need a particular skill level if one can quickly train up to doing the job. Faster human capital accumulation in a dimension tends to lower the correlation between initial skills and skill requirements in that dimension. Meanwhile, increasing κ_C, κ_M, ν_C, or ν_M tends to raise the correlation in the relevant dimension. It lowers the surplus for over and under-qualified workers and results in worker skills more closely aligned to the job requirements.[63] The NLSY79 cohort provides measures of initial skills and their skill requirements as described in Section 3.3. I target the observed correlation of initial cognitive and manual skills, $\mathbf{x}(0)$, with their respective job requirements, $\mathbf{y}$, for the cohort in the simulation that corresponds to the NLSY79. I include these correlations at years '79,

[62] I set the job-to-job transition rate target to 0.03 based on estimates in the literature (Moscarini and Thomsson, 2006).

[63] Formally, Lise and Postel-Vinay (2016) show that these parameters alter the set of jobs acceptable to each type of worker. The correlation serves as a metric to capture this information.

’81, ’84, ’87, ’90, and ’93, which constitutes twelve moments for these twelve parameters.[64]

Given $\mathcal{V}_0$ and all other parameters, the initial productivity parameters ($\alpha_{0,0}$, $\alpha_{C,0}$, $\alpha_{M,0}$, $\alpha_{CC,0}$, $\alpha_{MM,0}$) along with $\mathbf{x}$ shape parameters (ζ_C, ζ_M) govern wage differentials across $\mathbf{y}$ and occupational groups by extension. Fundamentally, information to obtain $\alpha_{C,0}$ comes from comparing wages of workers with similar $\mathbf{x}$ and y_M but different y_C (conditional on the model's other parameters). Information to obtain $\alpha_{CC,0}$ and $\alpha_{MM,0}$ comes from wage differentials of workers with different $\mathbf{x}$ but matched with the same $\mathbf{y}$.[65] Hence, I target initial average wages and wage dispersion for the high, medium, and low occupational groups described in Section 3.2. I also target decadal changes in average wages for these occupational groups as well as decadal changes at the 10th, 50th, and 90th wage percentiles. Changes at the 10-50-90 wage percentiles reflect changes in wage dispersion within occupational groups. These targets aims to capture the decadal trends in ($\alpha_{0,t}$, $\alpha_{C,t}$, $\alpha_{M,t}$, $\alpha_{CC,t}$, $\alpha_{MM,t}$). Additionally, I include the average mean and variance of log wages across the 1980s, 1990s, and 2000s as these levels contain further information on the α_t's and information on dispersion for the individual heterogeneity parameters (θ_0, θ_1). The use of decadal time trends for α_t's give twenty-four parameters for the thirty moments mentioned.

Finally, I target changes in the observed (equilibrium) distribution of skill requirements over time and decadal changes in employment shares across occupational groups. These targets identify the distribution of skill requirements, $\mathcal{F}_t(\mathbf{y})$, over the set of accepted $\mathbf{y}$'s (conditional on the rest of the model). I target the means, variances, and correlation of y_C and y_M in the initial year (1979) and their averages in the 1980s, 1990s, and 2000s to estimate the five parameters of $\mathcal{F}_t(\mathbf{y})$. I select the decadal change in employment shares across occupational groups, because this metric measures job polarization. The estimated $\mathcal{F}_t(\mathbf{y})$ must not only reproduce moments like mean and variance but also the preeminent feature of changes in the employment structure − job polarization. These targets yield twenty-nine moments for twenty parameters using decadal time trends for the scale, shape, and correlation parameters of $\mathcal{F}_t(\mathbf{y})$.

In summary, the parameters total sixty-four for eighty moments from 1979 to 2010. These moments consists of

1. decadal averages of mean and variance of log hourly wages
2. decadal averages of mean, standard deviation and correlation of (y_C, y_M) and in 1979
3. mean and standard deviation of wages within occupational groups in 1979
4. log change in occupational group employment shares and average wages over 1979-1989, 1989-2000 and 2000-2010 (Figure 5)
5. log change in 10-50-90 wage percentiles over 1979-1989, 1989-2000 and 2000-2010 (Figure 1)
6. average monthly job-to-job, employment-to-unemployment, and unemployment-to-employment transition rates over 1979-2010
7. average post-unemployment spell wage drop for the simulated NLSY79 cohort
8. differential between average wages and wages out of unemployment for the NLSY79 cohort

[64] As noted in Appendix B.5.1, I limit the NLSY panel to 1993, because sample attrition accelerates afterwards and makes the representativeness of the post-1993 sample suspect. In the data, I estimate $\mathrm{corr}(\widehat{x}_i(0), y_i)\ \forall i \in \{C, M\}$ rather than $\mathrm{corr}(x_i(0), y_i)$. I convert the model's $\mathbf{x}$ to $\widehat{\mathbf{x}}$ to compute the comparable model simulation target.

[65] A precise identification argument can restrict to workers out of unemployment or entering the labor force. These workers all possess the same bargaining power λ unlike workers with history dependent bargaining power.

9. correlations of $(x_C(0), y_C)$ and $(x_M(0), y_M)$ in 1979, 1981, 1984, 1987, 1990, and 1993 for the simulated NLSY79 cohort

10. average wages at ages 25, 35, 45, and 55.

5 Results

To present the results, I first show how well the model fits the data. Then I turn to what the parameter estimates say about the environment which yields this fit (e.g. how are skills valued relative to one another? how does this value change over time?). Next, I perform a series of decompositions to understand how and why the model fits. These decomposition shed light on the importance of the model's features (e.g. human capital accumulation, wage-setting employer competition). Finally, I look at what forces (e.g. routine-biased technological change) can explain the skill demand changes estimated with the model.

5.1 Model Fit

I estimate the model under three different assumptions and show how well each fits the data. The first two assumptions modify the horizon of foresight to shed light on the importance of anticipation when fitting the data. The last assumption modifies $\mathcal{V}_t(\mathbf{x})$ to inform on its importance for the model's fit. I take the first model (I) as the benchmark case. This benchmark considers the model with perfect foresight, human capital accumulation and decumulation, and the exogenous $\mathcal{V}_t(\mathbf{x})$ discussed in Section 3.3. Perfect foresight means agents know the entire path of z_t (i.e. $\mathcal{F}_t$ and f_t) following an initial shock starting at period 1. In model (II), I eliminate foresight from the benchmark. Agents do not anticipate changes to z_t and changes come as a surprise each period. A comparison of (I) and (II) grants insight into the importance of anticipation in reproducing the data. In model (III), I keep the perfect foresight benchmark, however I fix $\mathcal{V}_t(\mathbf{x})$ to $\mathcal{V}_0(\mathbf{x})$. This modification eliminates the reweighting of $\mathcal{V}_0(\mathbf{x})$ to obtain $\mathcal{V}_t(\mathbf{x})$. This reweighting adjusts for increasing educational levels and rising female labor force participation, holding fixed the within education-gender distribution of $\mathbf{x}$. A comparison of (I) and (III) informs on whether changes to $\mathcal{V}_0(\mathbf{x})$ help account for the data.

Figure 7 shows the models' fit to changes in employment shares and occupational average wages in the left and right panels, respectively. The left panel shows that the model replicates the job polarization observed in the data. Medium-skilled occupations shrank relative to both low and high-skilled occupations across decades. The right panel shows that the model mostly replicates changes in occupational average wages over the same period. Low and medium-skilled occupational wages fell while high-skilled wages rose on average in the 1980s. All wages rose in the 1990s with low and high-skilled wages rising more than medium-skilled wages (i.e. occupational wage polarization). The gap between occupational average wages expanded again the 2000s, however the model fails to match the fall in low-skilled occupational wages observed in the 2000s.[66] It also overestimates the increase in high-skilled occupational average wages. Overall, the model fits well to changes in employment and occupational wages and does not differ much over (I), (II), and (III). This outcome suggests neither anticipation effects nor changes in the distribution of skills supplied drive the broad occupational employment and wage patterns in the model.

Similarly, (I), (II), and (III) match the expansion of wage inequality across the wage distribution in the 1980s as shown in left panel of Figure 8 and Appendix Table 19. They also match expansion at the top of

[66] This fall occurs even when excluding the Great Recession 2007-2010.

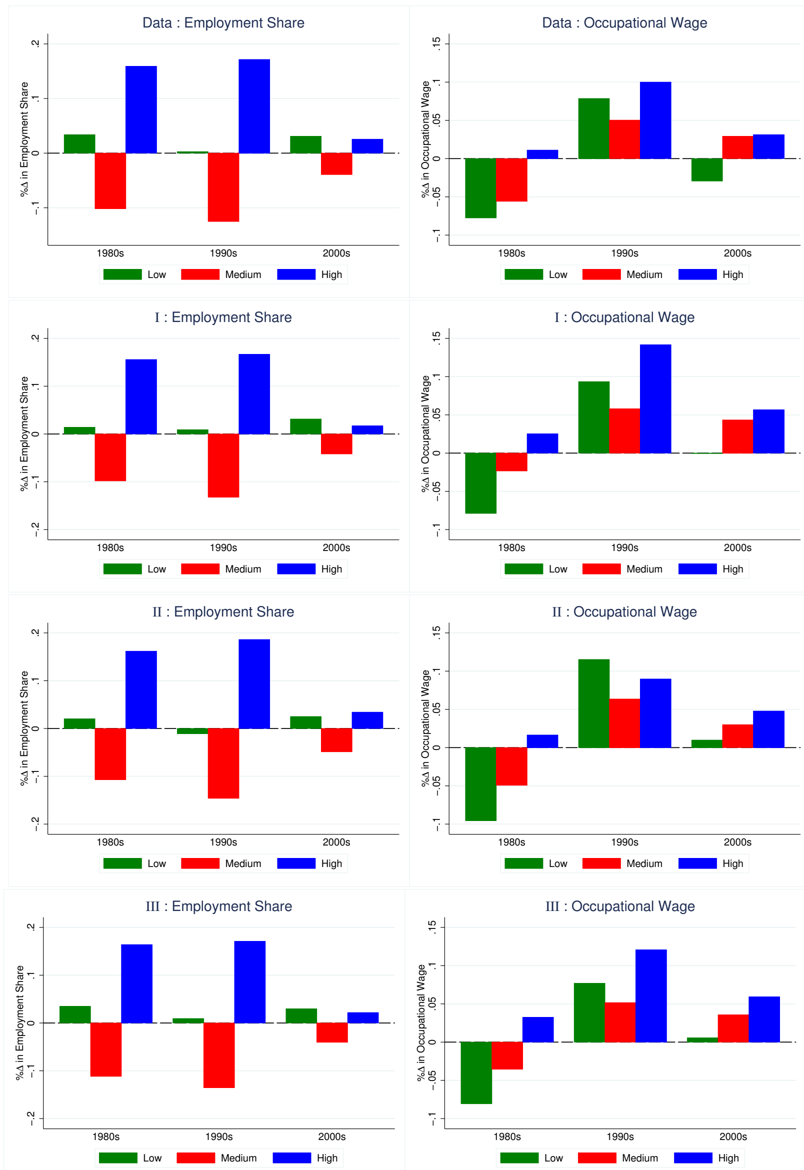


Figure 7: Employment Share (left) and Occupational Wage (right) Changes

the wage distribution in the 2000s. None generate the rise in wage in the lower tail of the distribution in the 2000s. However, the model does fit the 2000s after restricting to the period prior to the Great Recession (2000-2007) as shown in Figure 9. During the 1990s, the fit to changes at wage percentiles differs greatly over (I), (II), and (III). The perfect foresight benchmark (I) generates the right amount of growth at the 10th and 90th percentiles, but overestimates growth at the 50th percentile. This change translates to only minor compression (expansion) at the bottom (top) of the wage distribution compared to the dramatic U-shaped pattern in the data.

Eliminating foresight (II) worsens the fit as inequality expanded across the distribution in this case.[67] Thus, ignoring anticipation in this environment impedes matching the wage patterns we observe. In contrast, (III) produces the strong U-shaped change in the wage distribution. This outcome strongly suggests the adjustment to $\mathcal{V}_0(\mathbf{x})$ to construct $\mathcal{V}_t(\mathbf{x})$ affects the model's ability to produce wage polarization in the 1990s. This adjustment increases the share of college educated workers as well as female labor force participation to their observed levels over time, holding the within education-gender group distribution of $\mathbf{x}$ fixed. Comparing (I) to (III) shows holding the within education-gender distribution fixed worsens the model's fit. Ultimately, only wage polarization in the 1990s appears sensitive to the differing assumptions of (I), (II), and (III). Otherwise, the model fits wage distribution changes, occupational wage changes, and employment share changes well. Furthermore, the model matches occupational wages in 1979 and has a high correlation (above 0.96) with wage percentiles in 1979, 1989, 2000, and 2010 (Appendix Table 21). This correlation becomes particularly high (0.98) when excluding the extreme low (1-4) and high (96-100) percentiles. The model also tracks average wages and its increase closely as shown in Figure 10. It also tracks the increase in the wage dispersion despite overestimating wage dispersion. The job ladder effect causes the model to overestimate wage dispersion. Some workers take low wages in order to climb onto the job ladder, which creates a long left tail in the wage distribution (Appendix Figure 45).[68]

The model produces aggregate moments related to mobility and skill requirements. It also produces many but not all of the moments related to sorting and transitions to and from unemployment. The model generates moments from the equilibrium distribution of $\mathbf{y}$ shown in Figure 11, including the mean, dispersion, and correlation of skill requirements. However, it tends to underestimate the dispersion in cognitive skills and overestimate the correlation between manual and cognitive skills in the last decade. The model captures the correlation (i.e. degree of sorting) between initial cognitive and manual skills and their respective skill requirements (Table 4). But it fails to capture the size of the increase in the correlation of initial cognitive skills and cognitive skill requirements for the NLSY79 cohort (Appendix Figure 38).

The model generates the average monthly flow of employment to unemployment and vice versa as well as the average monthly job-to-job transition rate (Table 4). It overestimates the wage drop following an unemployment spell but roughly matches the average wage differential for wages out of unemployment compared to wage overall. The model also fits the targeted age-wage profile for the CPS (Figure 12) and the NLSY79 cohort (Appendix Figure 39). Overall, (III) delivers the best fit to all the target moments, explaining 95.4% of the variation in the target moments compared to 94.6% and 94.3% for (I) and (II), respectively (Appendix Table 17).[69]

[67] Workers at the 10th percentile in (I) became slightly more over-qualified in cognitive skills over 1989-2000. They obtained a larger increase in wages compared workers at the 10th percentile in (II) who become slightly more under-qualified in cognitive skills.

[68] Marginally qualified workers (where the surplus is just above zero) populate the lowest five percentiles in the wage distribution, increasing the left skewness. I later show the model with pure Nash Bargaining, which reduces the level of wage dispersion by eliminating the job ladder incentive to take low wage jobs.

[69] I discuss demographic heterogeneity in Appendix C.1.

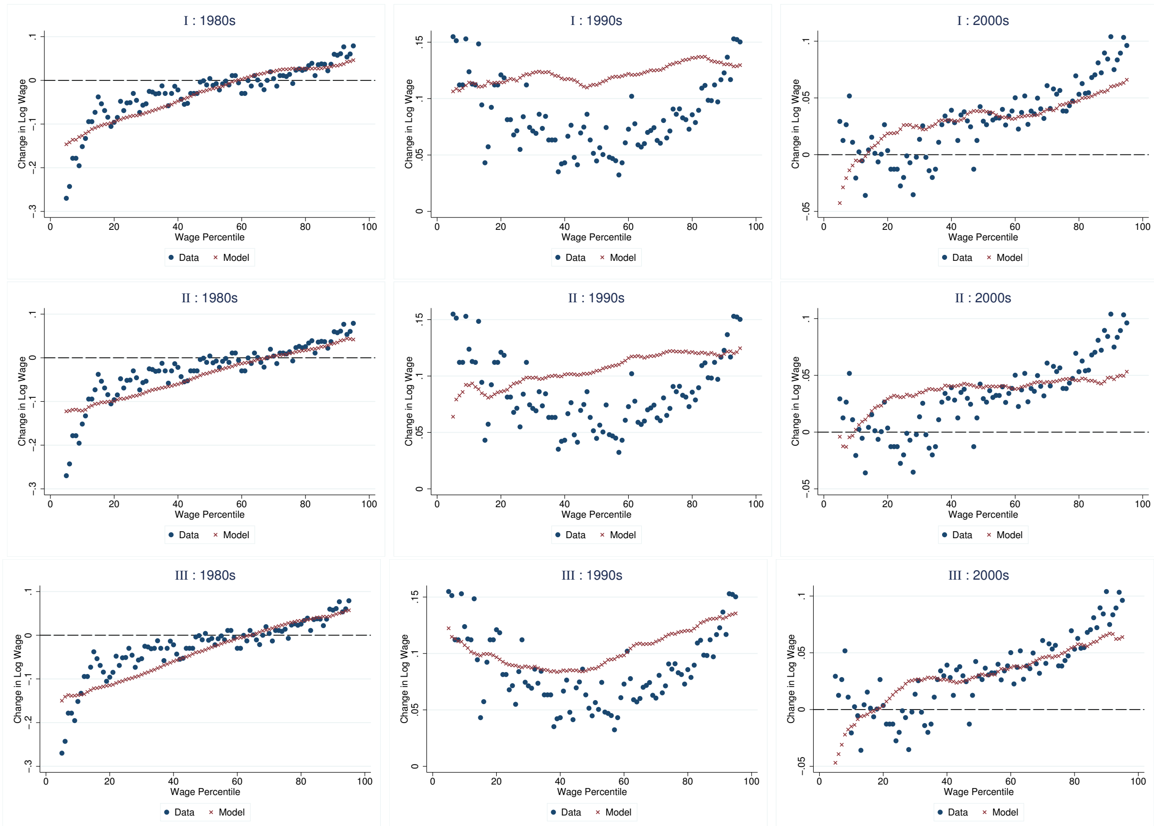


Figure 8: Wage Percentile Changes for I (top), II (middle) and III (bottom)

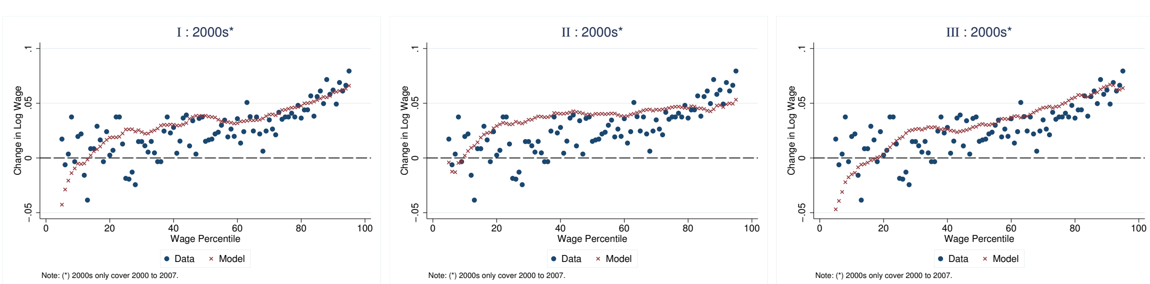


Figure 9: Wage Percentile Changes (2000-2007)

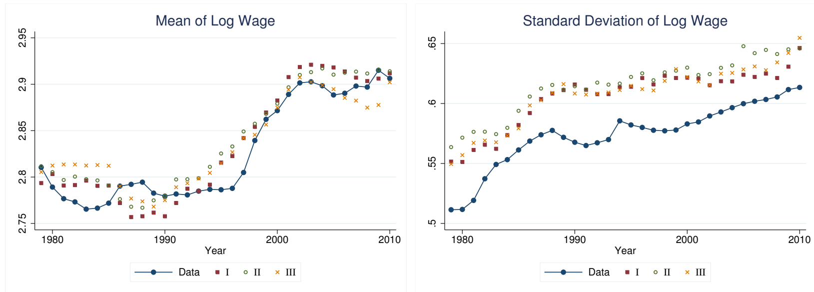


Figure 10: Mean (left) and Standard Deviation (right) of Log Wages

Table 4: Model Fit

	Data	I	II	III
Distribution of $\mathbf{x}(0)$ and $\mathbf{y}$				
$corr(x_c(0), y_c)$				
1980-1987	0.303	0.403	0.382	0.399
1988-1993	0.387	0.430	0.408	0.419
$corr(x_m(0), y_m)$				
1980-1987	0.078	0.083	0.064	0.065
1988-1993	0.083	0.053	0.050	0.040
Aggregate Job Flows				
Job-to-Job	0.032	0.024	0.035	0.032
Employment-to-Unemployment	0.015	0.016	0.015	0.017
Unemployment-to-Employment	0.261	0.266	0.277	0.262
Differential for U-to-E Wages (%)	-0.205	-0.234	-0.273	-0.243
Unemployment Spell Average Wage Drop (%)	-0.264	-0.430	-0.447	-0.417

5.1.1 Parameter Estimates

We now turn to understanding what features of the model yield this fit, starting with the model parameters. I present notable parameter estimates in Table 5 with the full set in Appendix Table 22. The parameter estimates indicate what environment appears consistent with the data. This environment consists of a higher valued, harder to accumulate skill and a lesser valued, easier to accumulate skill. The estimates also mirror the finding of Lise and Postel-Vinay (2016) that the model views different skills quite differently.[70] Mismatch between skills and skill requirements costs significantly more in the cognitive dimension in terms of output loss (governed by κ) and disutility of labor (governed by ν). Surplus lose due to under-qualification (i.e. $y_i > x_i$) remains higher than lose due to over-qualification in both skill dimensions. Cognitive skills accumulate much slower than manual skills. A worker learns manual skills fast and forgets them relatively slowly but learns cognitive skills slowly and forgets them relatively fast. These learning parameters (Γ_H, Γ_D) cause young workers to sort across jobs like prime age workers. Cognitive skill changes little over the life cycle

[70] It is worth noting that Lise and Postel-Vinay (2016) employ a different set of moments and data. They match their model solely to the longitudinal moments of the NLSY79 cohort using O*NET data. They use a plenthora of task content from O*NET to construct their scores for cognitive, manual and interpersonal skills. Here, I use mainly aggregate cross-sectional moments from the CPS, supplemented with information from the NLSY79 unavailable in the CPS.

Figure 11: Equilibrium Distribution of $\mathbf{y}$

in the model's estimation.[71] The estimates also indicate cross-skill complementarities in learning-by-doing with positive off-diagonal terms in Γ_H. For example, a worker possessing high cognitive skills can train up on-the-job to do more complex manual tasks faster than a worker with low cognitive skills and similar level of manual skills.

Production technology (f_t) shifts away from manual skills towards cognitive skills and from general skills to specific skills. Table 6 shows that general skills decline in their relative productive value (α_0) all else equal, biasing output towards specialized skills. Naturally, the model estimates cognitive skills to hold a higher baseline productive value (α_C) than manual skills (α_M), because workers in cognitive-intensive, high-skilled occupations earn more (Appendix Table 21).[72] However, production complementarities within tasks start on a comparable level but diverge over time. Call a worker with high x_C (x_M) a cognitive (manual) specialist. Table 6 shows that cognitive production complementarities (α_{CC}) increased twofold in the 1980s and continued to increase at a slower rate, benefiting cognitive specialists. Meanwhile, the relative productive value of manual specialists (i.e. α_{MM}) increased slightly in the 1980s with no notable increase afterwards. The change in distance between α_C and α_M pales in comparison to the change between α_{CC} and α_{MM}. Increased bias towards specific skills and divergence in the productive value of these skills characterize output

[71]Lower positive sorting across cognitive skills and requirements when young may come from yet-to-be known information about the worker's skill level rather than realized increases in cognitive skill levels over time. This process of learning about skills when young can result in more turnover and potentially less positive assortative matching across skill dimensions, e.g. Sanders (2012).

[72]Intuitively, the more difficult to acquire skill should be more valuable.

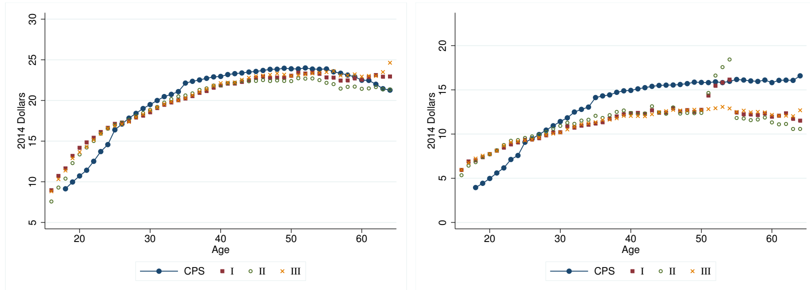


Figure 12: Mean (left) and Standard Deviation (right) Wage-Age Profile

Table 5: Time Invariant Parameters

	I	II	III
$\Gamma_H(1,1)$	0.0009	0.0045	0.0029
$\Gamma_H(1,2)$	0.0003	0.0020	0.0093
$\Gamma_H(2,1)$	0.0185	0.0014	0.0196
$\Gamma_H(2,2)$	0.0608	0.0525	0.0897
$\Gamma_D(1,1)$	-0.0148	-0.0113	-0.0209
$\Gamma_D(1,2)$	0.0000	0.0000	0.0000
$\Gamma_D(2,1)$	-0.0348	-0.0020	-0.0005
$\Gamma_D(2,2)$	-0.0331	-0.0535	-0.0330
ν_C	29.71	27.22	38.35
ν_M	14.19	0.0004	17.97
κ_C	130.8	103.0	128.7
κ_M	48.18	47.00	53.62

in the model.[73]

The distribution of skill requirements or skill demands exhibit a similar bias towards cognitive-intensive tasks. Figure 13 shows contour plots of changes in the density of $\mathcal{F}_t(\mathbf{y})$. It shows the job-polarizing changes in skill demands. Lighter areas show increased density while darker areas show decreased density. (I) and (III) estimate that the distribution of the skill demands concentrated more in the northwest quadrant and fell in the southeast quadrant. These changes to skill demands polarize employment as the model decompositions will show. Medium-skilled, manually-intensive jobs populate the southeast quadrant whereas high-skilled, cognitive-intensive jobs populate the northwest quadrant. (I) estimates a density increase spread across high-skilled and low-skilled jobs. In contrast, (III) concentrates in jobs in the high-skilled region of (y_M, y_C)-space. (II) estimates a proliferation of high and low-skilled jobs like (I). However, this proliferation occurs at all levels of y_C whereas (I) estimates a more concentrated change.

To summarize, the model fits the data well in many dimensions. It fits marginally better under full anticipation over no anticipation. It also fits better holding the skill endowment distribution fixed rather than adjusting it fully for between education and gender demographics. The model points to several key features

[73]Lindenlaub's (2017) assignment model also finds increasing cognitive complementarites and decreasing manual complementarities over the 1990s and 2000s. The model only has cognitive and manual skills.

Table 6: $f_t(\mathbf{x}, \mathbf{y})$ Parameters at Sample Dates

	I	II	III
$\alpha_{0,t=0}$	1.314	-8×10^{-5}	1.306
$\alpha_{0,t=121}$	-1.495	-1.905	-1.479
$\alpha_{0,t=267}$	-1.950	-1.542	-1.090
$\alpha_{0,t=384}$	-2.683	-1.208	-1.694
$\alpha_{C,t=0}$	20.26	19.56	19.23
$\alpha_{C,t=121}$	20.27	19.77	19.37
$\alpha_{C,t=267}$	19.80	18.18	19.36
$\alpha_{C,t=384}$	19.54	18.14	18.51
$\alpha_{M,t=0}$	-0.775	1.247	-1.283
$\alpha_{M,t=121}$	-0.853	0.646	-1.383
$\alpha_{M,t=267}$	-0.516	0.571	-1.383
$\alpha_{M,t=384}$	0.344	0.282	-0.379
$\alpha_{CC,t=0}$	9.914	10.62	8.379
$\alpha_{CC,t=121}$	21.23	16.62	21.01
$\alpha_{CC,t=267}$	31.83	24.52	32.68
$\alpha_{CC,t=384}$	34.48	28.04	36.58
$\alpha_{MM,t=0}$	8.387	8.877	8.615
$\alpha_{MM,t=121}$	9.055	10.14	10.46
$\alpha_{MM,t=267}$	6.261	6.174	8.193
$\alpha_{MM,t=384}$	5.930	2.733	7.347

to fit the data. First, cognitive skills accumulate slower and decline faster (relative to their accumulation speed) than manual skills. Second, cognitive skills hold higher productive value than manual skills, and this value increased over time to favor cognitive specialists and slowed more recently. Third, the distribution of skill demands exhibits polarization. The first feature says some skills must be slower to adjust to understand the data. I assess the importance of this feature in the model decompositions. The last two features come as no surprise. Education strongly correlates with cognitive skills (Figure 6), and wage returns to education have become more convex over the period under consideration.[74] The lens of the model says the "convexification" of the returns to education reflect changes to the productive value of cognitive skills. Specifically, cognitive production complementarities increased (at a decreasing rate) since the 1980s. Of course, the distribution of available jobs, $\mathcal{F}_t(\mathbf{y})$, affects the allocation of workers. This allocation exhibits polarization, so the distribution of skill demands polarized as expected.

5.1.2 Skill-Biased Technical Change v. Task-Biased Technical Change

The model provides alternative interpretations to skill-biased technical change and task-biased technical change. Skill biased technical change conceives of a labor market consisting of high and low skilled workers. Technological progress increases the productivity of high-skilled workers and wage inequality expands as a result (Acemoglu and Autor, 2011). Task-biased technical change conceives of a labor market consisting of a

[74]See Valletta (2016) for a detailed discussion on the "convexification" of the returns to education.

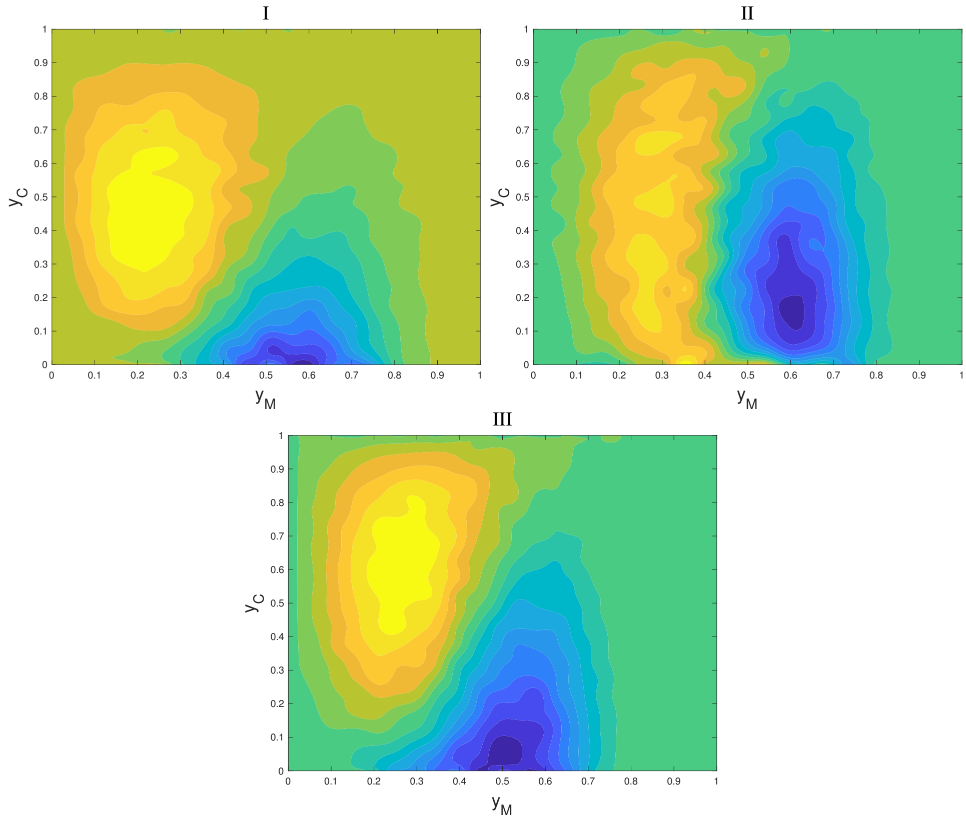


Figure 13: $\mathcal{F}_{385}(\mathbf{y}) - \mathcal{F}_1(\mathbf{y})$

mix of tasks, and workers use their skills to do said tasks. The returns to performing a specific type of task increase and wage inequality may expand, contract, or do both but in different parts of the wage distribution (Acemoglu and Autor, 2011; Lindenlaub, 2017). The outcome depends on which workers reallocate to which tasks. Routine-biased technical change serves as a notable example of task-biased technical change. Routine-biased technical change lowers the relative value of medium-level skills used to do routine tasks like assembly or clerical work. Workers select out of these medium-wage tasks as their relative value falls. This selection produces an expansion above the median wage and expansion or contraction below it depending on which workers move into low or high-skill tasks. Some papers consider the 1980s to represent skill-biased technical change while the 1990s represent more task-biased technical change.[75]

The model conveys skill-biased technical change in the 1980s as productivity changes favored specific skills (x_C, x_M) over general skills (x_G). Estimates of the baseline return to general skills (α_0) drop in the 1980s, while production complementarities $(\alpha_{CC}, \alpha_{MM})$ rise. The model conveys task-biased technical change as a productivity shift towards cognitive skills over manual skills. In the 1990s, the fall in α_0 stops or decelerates while α_{MM} begins to fall or stagnate and α_{CC} continues to rise. Thus, productivity estimates move towards move towards cognitive skills away from manual skills. Hence, skill-biased technical change consists of specialization (i.e. a shifts towards specific skills). Meanwhile, task-biased technical change consists of shifts towards a particular specific skill and away from another. In this sense, the model exhibits skill-biased technical change in the 1980s and task-biased technical change in the 1990s.[76]

[75] See Acemoglu and Autor (2011), (Boehm, 2017), or Mishel, Schmitt, and Shierholz (2013) for further discussion.

[76] Alternatively, we can interpret the estimates through definition of Lindenlaub (2017). In this case, there is only task-biased technical change as increased (α_C, α_{CC}) convey skill-biased and task-biased technical change, respectively.

5.2 Decompositions

Examining the model parameters grants broad insight into what assumptions and parameters allow the model to fit the data. However, they do not readily explain what factors and assumptions drive results like how (III) best fits polarization in the 1990s while wages in (II) do not polarize at all. To this end, I use a series of decompositions to unpack the role of the exogenous factors and model features. First, I comparatively examine the model versions shown to gain insight into the role of anticipation and the skill supply distribution, $\mathcal{V}_t(\mathbf{x})$. Then I perform decompositions to evaluate the importance of human capital evolution (IV), heterogeneous specific human capital (V), changes in the distribution of skill demands (VI), changes in production technology (VII), multidimensional skills (VIII), and employer wage-setting competition (IX).

For each decomposition, I simplify the benchmark model (I) in an aspect and re-estimate the model to measure the contribution of the relevant factor. Counterfactual analysis alone is unsuitable to measure the contribution of a factor. The job selection and wage setting mechanisms in the model interact with all of these factors. They amplify or dampen their effects even in partial equilibrium. For example, suppose we want to measure the importance of changes in the distribution of skill demands in matching the data. The relevant counterfactual should measure how much of the data we account for in the absence of changes to the distribution of skill demand. The gap between this measure and what we account for when changing $\mathcal{F}_t(\mathbf{x}, \mathbf{y})$ measures its importance. A naive counterfactual holds $\mathcal{F}_t(\mathbf{x}, \mathbf{y})$ fixed without adjusting the model parameters. This counterfactual ignores the increased importance of $f_t(\mathbf{x}, \mathbf{y})$ in the allocation of workers to jobs in the absence of changes to $\mathcal{F}_t(\mathbf{x}, \mathbf{y})$, thereby overstating the importance of $\mathcal{F}_t(\mathbf{x}, \mathbf{y})$. I begin the decompositions comparing (I), (II), and (III) as they form the basis as to why I perform the decompositions that follow. How well they match wage polarization in the 1990s distinguishes these versions of the model.

5.2.1 Wage Polarization

How do wages polarize in the 1990s? During this period, the distribution of cognitive skill requirements shifts towards cognitive skills, and production complementarities in the cognitive dimension increase. The opposite occurs in the manual skill dimension. Thus, the parameters suggest a task-specific relative demand shift towards cognitive skills away from manual skills in the 1990s in contrast to a shift towards both specific skills away from general skills in the 1980s. This shift results in a proliferation of cognitive jobs and an increase in their average wages (i.e. occupational upgrading). Meanwhile, a large deceleration in loses due to specialization (i.e. α_0 stabilizes or increases) and an increased level of general skills (e.g. older workers) drive wage gains for workers in the low-skilled occupation.[77] Thus, the model produces polarization in occupational wages and employment similar to the data in cases (I), (II), and (III).

However, neither job polarization nor occupational wage polarization serve as necessary or sufficient conditions to generate wage polarization.[78] The model allows us to clarify how wage polarization occurs. We observe similar changes to the distribution of skill demands and productivity parameters across (I), (II), and (III).[79] We observe polarization in the average wage in each occupational group. Moreover, we observe similar trends in changes in the wage distribution in the 1980s and 2000s. Yet only (I) and (III) lead to inequality expansion above the median wage and compression below the median.[80] Only (III) produces the

[77]Recall that general skills amplify the output of specific skills. This feature generates the lifecycle profile of wages. Alternatively, we can interpret α_0 as reflecting the effect of economic growth on wages rather than skill specialization. However this model is not a growth model, leaving such an interpretation ambivalent.

[78]Boehm (2017) proves this claim theoretically in a static, competitive Roy model.

[79]Correlation for all estimates are 0.990 for (I) and (II), 0.997 for (I) and (III), and 0.989 for (II) and (III).

[80]It is possible to eliminate inequality expansion below the median in (II). However, it increases low-skilled wage growth well above its target value (Appendix Figure 46) and thus is not the optimal fit to the data.

dramatic U-shaped polarization which occurs in the data. Comparing these versions of the model grants insight into how wage polarization in the 1990s arose.

Table 7: Change in Wage Percentile for Median Worker

Occupational Group	Data	I	II	III
High	+1	+0	−3	+0
Medium	−1	−4	−3	−3
Low	+1	−1	+2	−1

In the 1990s, marginal expansion occurs below the median in (II), because wage growth rose disproportionately in the low-skilled occupational group in (II). Some workers in this group overtook workers in the (shrinking) medium occupational group.[81] Consequently, the lowest percentiles in 2000 do not reflect their wage gains. Table 7 shows the change in the wage percentile for the median worker in each occupation group. Low-skilled occupation workers gain the most on medium-skilled workers in (II). In contrast, (I) produces some wage compression in the bottom half without disproportionately increasing wages in the low-skilled group. In (I), workers and employers anticipate technological and skill demand changes. They agree to a wage schedule that in part backloads these expected gains to incentivize the worker to stay at the job. Anticipation in (I) puts downward pressure on wages where workers and employers expect gains, reducing this overtaking effect while still allowing these wages to rise. In fact, removing foresight under the estimates of (I) results in inequality expansion across the wage distribution in the 1990s and much more extreme occupational and wage changes (Appendix Figure 47). In contrast, workers and employers do not anticipate such gains and loses in (II). Adding foresight under the estimates of (II) causes a negligible inequality contraction across the entire wage distribution (Appendix Figure 48). Overall, the benchmark foresight model (I) fits marginally better to the data as it spreads out of gains over time. However, the improvement over the model with no foresight (II) remains small.

Model (III) constitutes a marked improvement over (I) with respect to wage polarization in the 1990s. Recall that (III) estimates the model without adjusting $\mathcal{V}_t(\mathbf{x})$ over time. The adjustment reweights the within education-gender distribution of $\mathcal{V}_0(\mathbf{x})$ to match their demographic shares in the labor market. This improvement suggests the reweight over adjusts of the distribution of skill endowments, because the distribution of skills within these groups changed over time. Notably, women exhibited a lower mean for manual skills in the NLSY79. The NLSY97 confirms a within gender upward (downward) shift in documented manual skills for women (men) (Appendix Figure 50). The distribution for men and women look similar for cognitive skills in the NLSY79 (Figure 6) but appear to align on manual skills over time. Thus, the adjustment for rising female labor force participation re-enforces a gender bias in documented manual skills which diminished over time. Thus, we see the shape of the distribution of skill endowments matters greatly to produce wage polarization in this model. Holding the share of female labor force participation fixed but adjusting for the rising share of college education workers (X) generates more of the U-shape change to wages in the 1990s (Figure 20). Holding $\mathcal{V}_t(\mathbf{x})$ fixed does not mean the distribution of skills remains fixed. Human capital accumulates and decumulates over the life cycle and in response to structural change. Both shape the endogenous skill distribution in the labor market. In both (III) and (X), manual skills accumulate faster compared to the benchmark (Appendix Table 22).

[81] In 1989, about 16% of low-skilled occupation worker earned less than the 10th percentile middle-skilled occupation worker.

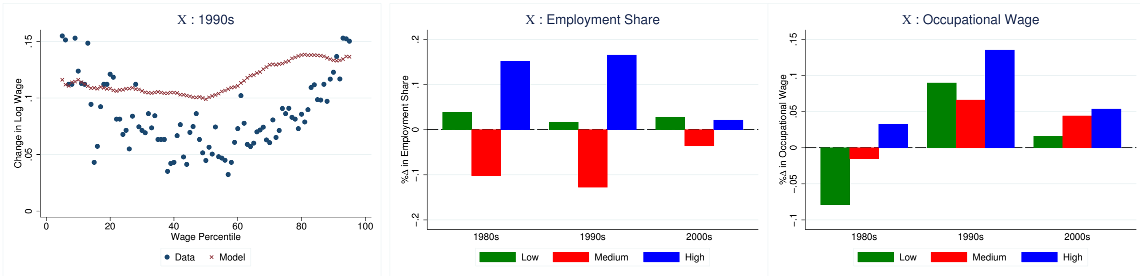


Figure 14: Fixed Female Labor Force Participation

Comparing (I) and (III) clarifies why the skill endowment distribution contributes to a more dramatic U-shape change in wages. The skill endowment distribution in (I) results in too many workers in the medium occupation concentrated in the lower tail of the wage distribution. Figure 15 shows the (smoothed) employment share for the medium-skilled occupation at every wage percentile. This curve shows workers in medium-skilled occupation remain prolific in the middle of the wage distribution and less so in the upper and lower tails. The curve shifts downward as the medium-skilled employment share shrinks. It also becomes "less concave" at lower percentiles from 1989 (solid red) to 2000 (dashed green), meaning workers in the medium occupations concentrate more in lower wage percentiles. In (I), these workers start out more prolific in the lower wage percentiles in 1989 and move downward. Consequently, more low-skilled occupation workers overtake them in the wage distribution between the 10th and 50th percentiles. High-skilled occupation workers move into the 50th percentile as the medium-skilled occupation shrinks and medium occupation workers move down the wage distribution, driving up wages at the 50th percentile. Thus, the movement of

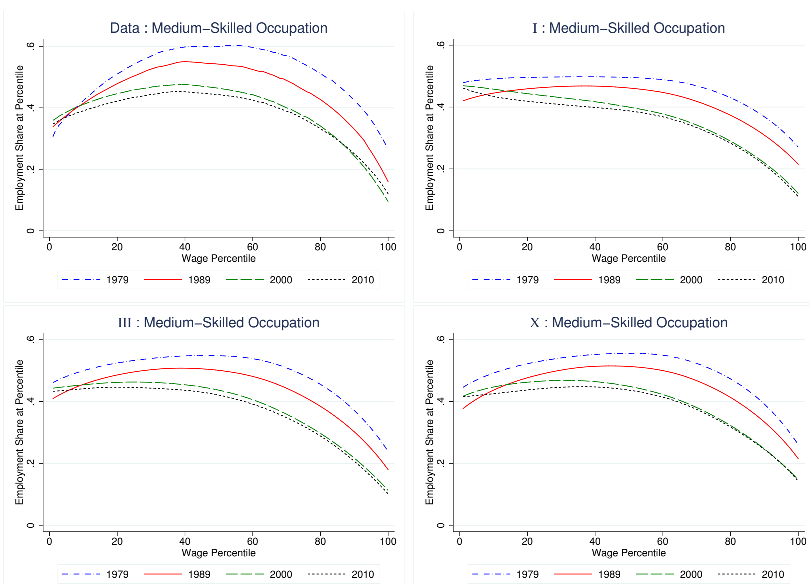


Figure 15: Employment Share at Wage Percentiles

medium occupation workers downward causes the model to overestimate the increase in wages at the 50th percentile. This overestimation worsens in case (I) where these workers start out more concentrated in the lower percentiles compared to (III) and (X).

By 2000, this percentage fell by 2.36 percentage points (ppts) in (III), 2.66 ppts in (I), and 4.21 ppts in (II).

The endowment distribution of manual skills accounts for more medium occupation workers in lower wage percentiles in (I). The adjustment to construct $\mathcal{V}_t(\mathbf{x})$ skews the manual skills distribution negatively. This results in more manual skill under-qualification among medium-skill workers and pushes them to lower wage percentiles.[82] Of course, manual skills accumulate toward the job requirements, but matching the positive correlation between initial manual skills and manual job requirements constrains how fast manual skills can accumulate.[83] Thus, the skill endowment distribution is crucial to matching wage polarization in the 1990s in this model. This result contrasts starkly with Lindenlaub (2017) who concludes changes in the distribution of skill endowments are not crucial to account for wage polarization.

5.2.2 Learning Frictions

Time-consuming human capital evolution and heterogeneous specific human capital cause under and over-qualification arise in this model. Call them learning frictions or matching frictions. They result in imperfect matches (i.e. $\mathbf{x} \neq \mathbf{y}$). How important are such frictions to accounting for wage and occupational changes from the 1980s to 2000s? I perform two decompositions to answer this question. The first removes specific human capital accumulation and decumulation from the benchmark model. This decomposition evaluates the explanatory power of learning on-the-job. The second removes the matching friction in the model so that human capital changes instantaneously. This modification equates to making cognitive and manual specific human capital homogeneous where $\mathbf{y}$ serves as a permanent, match-specific productivity shock. This decomposition evaluates the importance of misalignment between skills and skill requirements in account for occupational and wage changes.

Eliminating specific human capital accumulation and decumulation makes little difference to the overall model fit relative to the benchmark.[84] More under-qualified medium occupation workers end up in the lowest wage percentiles due to their inability to acquire more manual skills. Consequently, workers in the low-skilled occupation overtake them, making wage polarization difficult to generate compared to the benchmark. However, the model fits just as well on occupational employment and wage changes. This outcome suggests limited importance for skill loss and acquisition relative to factors like structural change in $\mathcal{F}_t$ and f_t. Of course, this indication only says specific human capital evolution remains of limited importance to reconcile broad wage and occupational changes. Specific human capital accumulation may be crucial to understand a wide set of phenomena like job promotion paths.[85] Also, the acquisition of general human capital over the life cycle remains important to matching growth at the lower percentiles.

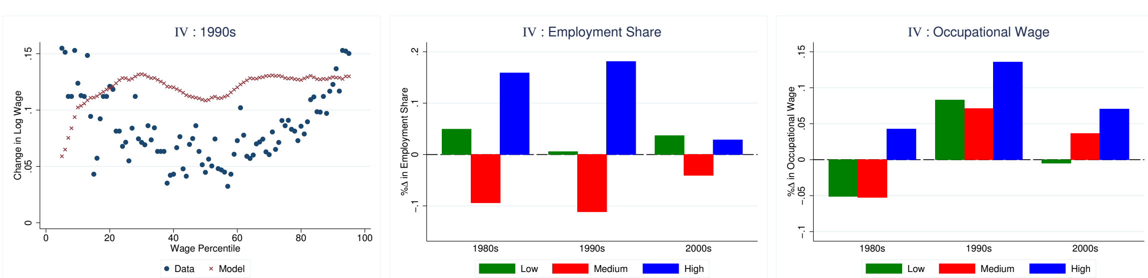


Figure 16: Fixed Specific Human Capital

[82]The average difference between x_M and y_M in 1989 is -0.080,0.001, and -0.006 for (I), (III), and (X), respectively.
[83]Faster accumulation drives this correlation down as a worker may obtain a job with possessing a level of manual skill well below what is required.
[84]Appendix C displays full results for all decompositions.
[85]See Gibbons and Waldman (2004).

The next decomposition eliminates the matching friction caused by heterogeneous specific skills. Consequently, it also eliminates the concepts of under-qualification, over-qualification, and sorting. Permanent $i.i.d.$ match-specific productivity shocks constitute $\mathcal{F}_t(\mathbf{y})$ and each worker offers an indivisible, homogeneous unit of cognitive and manual skill. The α_t's determine aggregate productivity, while $\mathcal{F}_t(\mathbf{y})$ determines the idiosyncratic productivity of the specific skills. This model turns out to fit the data well with respect to wage levels, wage dispersion, and occupational wage and employment changes (Figure 16, Appendix Table 25, Appendix Figure 52). It accounts for just under half (45%) of the variation in the target moments that the benchmark model explains (95%). Therefore, heterogeneous specific human capital and the skill mismatch and sorting it produces account for the other half of the model. $\mathcal{F}_t(\mathbf{y})$ permits the model to generate job polarization, while α_t's produce occupational wage expansion and polarization. The model also generates the observed wage expansion in the 1980s and 2000s, but it does not generate wage polarization. It fails to match occupational wages for the middle-skilled group (last column of Table 8), resulting in large shares of medium occupation workers at the bottom of the wage distribution. Occupational wage polarization pushes these workers further down. Wage polarization cannot occur as a result. Clearly, heterogeneous specific human capital (along with the skill mismatch and imperfect sorting it generates) appears crucial to reconcile wage and occupational changes from the 1980s to 2000s. Overall, the model suggests matching frictions matter greatly while skill evolution appears non-essential for this matter.

Table 8: Mean Occupational Wage in 1979 (V)

	Data	I	II	III	V
High	25.344	25.532	25.311	25.023	27.292
Medium	18.216	17.715	17.967	17.855	14.858
Low	14.410	15.126	14.411	15.106	14.072

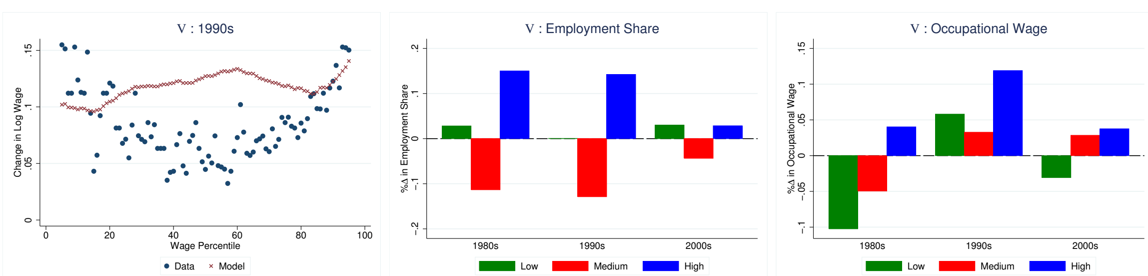


Figure 17: Homogeneous Specific Human Capital

5.2.3 Productivity and Skill Requirements

The controversy surrounding the job-polarization explanation for wages centers on continuous job polarization but discontinuous wage trends. Mishel, Schmitt, and Shierholz (2013) argue that the long-run secular trend of job polarization cannot account for the reversal in wage expansion below the median in the 1990s. The previous decompositions indicate changes in $\mathcal{F}_t$ and f_t can account for a large portion of the data even without heterogeneous task specific capital. What individual role do these structural changes play in shaping consistent job polarization but inconsistent wage polarization? Can they reconcile these seemingly contradictory phenomena? I decompose the model first holding $\mathcal{F}_t$ and then f_t fixed to shed light on how each shape occupational and wage changes. The outcome indicates that the job-polarizing distribution of skill demands acts as a force of wage compression across occupational groups and the wage distribution broadly. Whereas, changes in productivity levels act as a force of inequality expansion between occupational groups and all across the wage distribution. Whichever dominates governs whether we observe inequality growth or wage polarization.

Figures 18 and 19 show the occupational and wage results re-estimating the model holding either $\mathcal{F}_t$ or f_t fixed, allowing the other to evolve. f_t alone (Figure 18) generates inequality expansion in the 1980s and 2000s to a lesser extent. It also delivers wage expansion across occupational wages in the 1980s. However, it fails to generate enough sorting to match job polarization in any period. It fails to even match the patterns, let alone the magnitudes. It also predicts wage expansion across occupations in the 1990s but wage contraction across the entire distribution.[86] The model again indicates specialization and its deceleration but fails to estimate the extent of task-biased technical change. Thus, changes in the distribution of skill demands help identify task-biased technical change to account growth at the 90th percentile and job polarization. $\mathcal{F}_t$ alone (Figure 19) produces general patterns capturing job polarization, however it fails to generate inequality expansion across any decade. In fact, wages compress in all three decades as medium-skilled workers upgrade to the high-skilled occupation (i.e. occupational upgrading). Such changes in the distribution of skill demands offset the inequality expanding force of productivity shifts. This decomposition reveals $\mathcal{F}_t$ and f_t counteract to produce a consistent pattern of job polarization with varying changes to the wage distribution over time. Quantitatively, they appear equally important when comparing their overall fit to the data.[87]

5.2.4 Multidimensional Skills

Acemoglu and Autor (2011) suggests a task-based framework with at least three skill groups best serves to analyze job polarization and wage changes. Generating wage polarization requires at least two. Many models of labor market sorting reduce human capital to a single index to evaluate the impact of technological change, e.g. Kantenga and Law (2017). Evidence like Kambourov and Manovskii (2002) points towards occupation-specific rather than task-specific skills. Thus, we might expect tasks to miss out on important differences between occupations. For instance, a ballerina cannot smoothly transition to being a glass cutter even though both require high levels of manual dexterity and moderate levels of cognitive ability. This begs the question: how important is the occupational heterogeneity which tasks fail to capture? More pointedly, what is the sufficient number of tasks/skill required to reconcile occupational and wage changes? This paper focuses on major occupation groups, but the model has any number of occupations with three skills and only two tasks. I address this question in the most direct way possible by "eliminating" one of the tasks from the model and comparing it to the case with two tasks.

The multidimensional nature of specific tasks define occupational groups. I aim to preserve this definition to make this two skill, two task model comparable. To do so, I relegate manual tasks and skills to an entirely descriptive role, setting $\alpha_{M,t}$, $\alpha_{MM,t}$, ν_M, and κ_M all to zero. In this case, manual skills and tasks play no role in job selection or wage setting. Manual tasks merely define which occupation we call high, medium, and low. This change preserves comparability while effectively eliminating manual skills/tasks from the model mechanisms. The model effectively consists of cognitive specific human capital, general skills, and a cognitive

[86] Lindenlaub (2017) matches wage polarization with a static assignment model, however that model imposes no consistency with respect to neither occupational wage changes nor employment shares.

[87] This result again starkly contrast with the conclusion of Lindenlaub (2017) that changes production complementarities outweigh the importance of changes in the distribution of skill requirements.

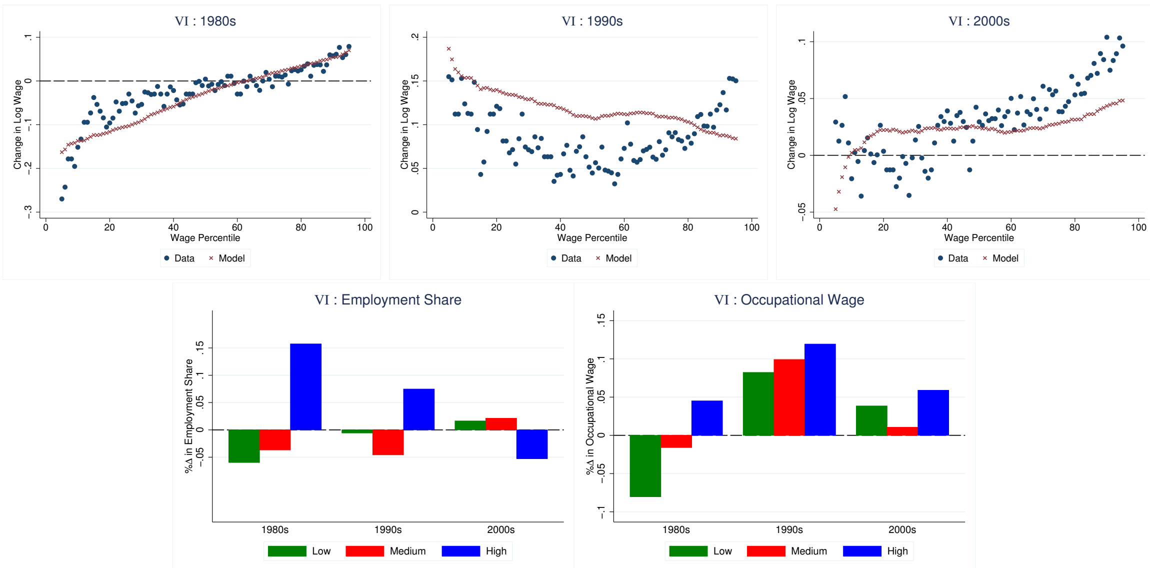


Figure 18: Fixed $\mathcal{F}(\mathbf{y})$

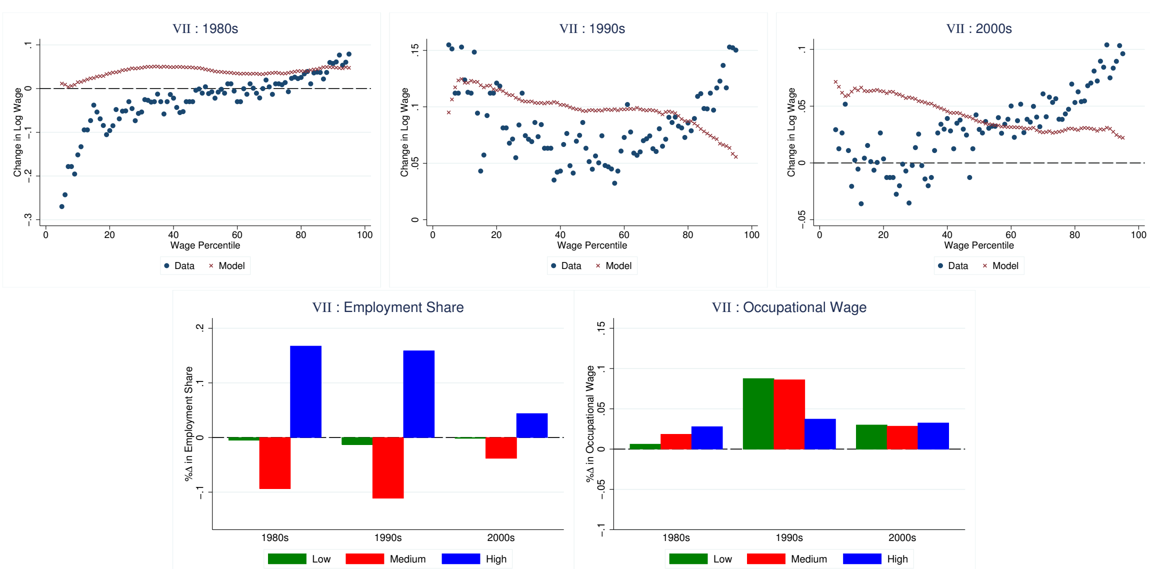


Figure 19: Fixed $f(\mathbf{x}, \mathbf{y})$

task.

This decomposition produces a striking result. It fits wage polarization better than the benchmark and explains 88% of the variation in moments overall. It also fits wage polarization just as well as (III) (Figure 20). Merely assigning different labels over time produces the patterns of job polarization and occupational wage expansion/polarization, although not the magnitudes. Hence, we see that cognitive skills/tasks and general skills provide enough content to reconcile patterns of wage changes and changes in the occupational wage structure. The model even replicates moments from the marginal distribution of y_M in matching occupational patterns (Appendix Table 29). Though this lens, specialization in cognitive tasks account for all of these patterns. Unsurprisingly, this model (VIII) cannot replicate the level of correlation between cognitive and manual task complexity (y_M), because manual complexity (y_M) serves as a mere label. However, it does decline as it does in the data. This model improves over the benchmark in wage polarization for

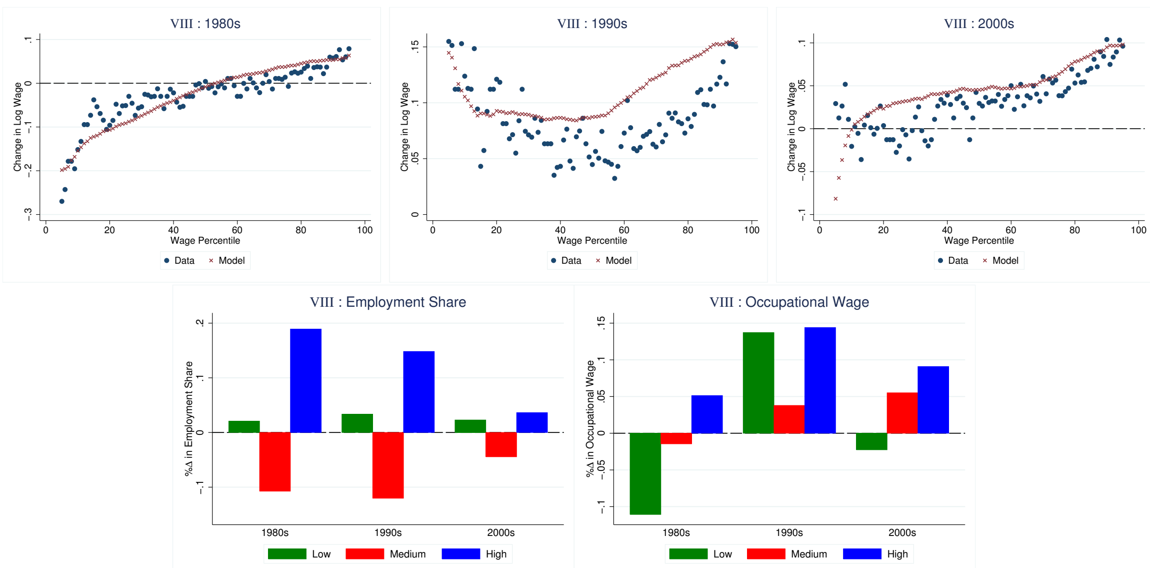


Figure 20: Non-Productive Manual Skills ($\alpha_M = 0,\ \alpha_{MM} = 0,\ \nu_M = 0,\ \kappa_M = 0$)

similar reasons to (III). It eliminates under-qualification in manual skills and thus concentrates more medium occupation workers in middle wage percentiles. This "cognitive-biased" technical change does not push these medium occupation workers down the wage distribution like task-biased technical change. Increases in production complementarities increase convexity in the upper half of the wage distribution. Meanwhile, a deceleration in the fall of α_0 and higher x_g raise wages disproportionately at the bottom. The lens of this model says the move from inequality expansion to wage polarization comes from a slowdown in specialization rather than a different type of technological change. The shift towards higher cognitive task complexity produces some tasks that require no specialized skills. For example, suppose a technology firm expands to a new location. This new building raises demand for local low-skilled protective services and food service workers.

On one hand, this outcome suggests no need to model a large array of tasks and skills to reconcile the patterns we observe. Idiosyncrasies in task content do not prevent us from understanding broad changes using task complexity alone. It appears enough that tasks vary in their cognitive complexity to account for changes in the wage distribution. Identifying occupations by applying the labels constructed in the data yield patterns for broad occupational groups. The presence of two skills – not three – and one task provide enough information to generate these patterns in a frictional setting with heterogeneous human capital. The inability to reconcile wage polarization with a "canonical" competitive model inspired use of the tasks framework (Acemoglu and Autor, 2011). Frictions (like search and learning frictions) provide an alternative, tractable way to enrich the environment instead of expanding the dimensionality of tasks to capture an intractable, large set of occupations. On the other hand, this simpler model remains unsatisfactory. It seems unreasonable that non-cognitive skills hold such little value in rewarding productivity or allocating jobs. We can interpret it as the value of cognitive skills and age encompass the value of all other non-cognitive skills. Data suggest otherwise as these skills/tasks do not correlate perfectly. It also appears skills like interpersonal skills hold some importance for wages and job allocation distinct from cognitive skills (Jaimovich, Siu, and Cortes, 2017).

5.2.5 Stationarity vs. Trends

How much of the same conclusions do we draw when conceptualizing technological change as a one-time, permanent shock rather than ongoing structural change? The following decomposition determines the significance of looking at a transition path to examine occupational structural change. The benchmark model imposes discipline in the labor market across time. The distribution of skill requirements and productivity evolve gradually, some parameters remain fixed, and cross-sectional outcomes aggregate from overlapping cohorts. A non-stationary (partial) equilibrium transition path maps out the labor market from 1979 to 2010. Many papers take an alternative approach to technological change or demand shifts. Instead, they consider them as a one time, permanent adjustment and estimate a series of steady state models over sub-periods.[88] They then use estimates over each sub-period to perform counterfactual analysis and make inference about technological change.

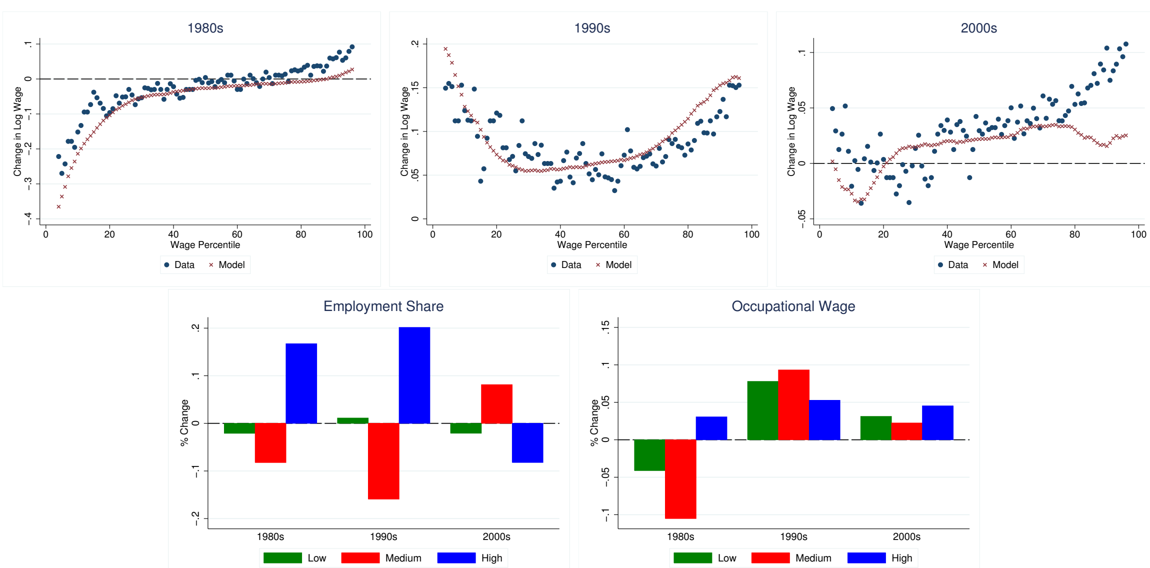


Figure 21: Repeated Stationary Model

I estimate a stationary version of the model over year long sub-periods. This version eliminates human capital evolution and foresight over structural changes. I estimate the stationary model to match annual

[88]e.g. Lindenlaub (2017); Kantenga and Law (2017). The question arises as to how long a period makes a steady state.

levels of the target moments when available.[89] This model generates strong U-shaped wage polarization in the 1990s. In doing so, it fails to match wage and occupational changes otherwise (Figure 21). In fact, it overestimates wage polarization in the 1990s (Appendix Table 33). Imposing consistency over time greatly improves the fit to occupational wage and employment changes. In this frictional setting, the accumulation of decisions in an ongoing transition appear to better describe occupational structural change than two entirely different states of the world.

5.2.6 Employer Competition

Lastly, I re-estimate the model with pure Nash Bargaining to see if employer competition made any difference to the model fit. Nash Bargaining does not change the allocation of workers to jobs under the same parameters as (I) or (III), because it only affects the *split* of the surplus, not the surplus itself. However, it does affect wages as shown in Appendix A.4. Nash Bargaining results in a marginally worse fit (Appendix Table 31, Appendix Figure 53), but ultimately employer competition makes little difference compared to Nash.

5.3 External Validation

I show the model can evaluate hypotheses and produce findings in the literature to support to its validity. The model yields ambiguous predictions as discussed in Section 2. This lack of prediction makes the model flexible enough to match wage and occupational patterns over time. However, it also makes the model difficult to validate. Natural questions for such a quantitative model include: how plausible are the model's insights? What can we observe in the data to evaluate (if not test) the model's validity? For example, the model indicates workers hold higher general skills in low-skilled occupations in 2000 due to aging. Thus, we ought to observe older workers in low-skilled occupations in 2000 compared to 1989. If workers in the low occupations become younger on average, then we might question what the model says about wage polarization. Average ages increase in this occupational group in the CPS data, which is consistent with the model.

Table 9: Task Price Polarization (in Log Points)

	Boehm (2017)	I	II	III
$\Delta(\pi_A - \pi_R)$	27.3	8.4	48.3	34.9
$\Delta\pi_R$	-5.2	-0.0	-17.4	-19.2
$\Delta(\pi_M - \pi_R)$	32.0	4.6	5.6	15.0

A good model replicates at least some relevant findings in reduced form approaches. In contrast to this model, the competitive framework offers stronger predictions. Notably, task-biased technical change results in polarized "task prices" if nothing else (Boehm, 2017). I estimate these competitive "task prices" as a reduced-form validation exercise. Suppose we conceive of wages in the general terms of a competitive Roy-style assignment model. Wages equal the sum of skill prices times the worker's skill level. Based on this framework, Boehm (2017) develops a reduced-form method comparing NLSY cohorts to estimate relative changes in task prices for manual (π_M), routine (π_R), and abstract (π_A) tasks. He finds evidence of "task

[89]Correlations between initial specific skills and current job requirements are unavailable in 2000 and 2010 due to imposed data restrictions. Instead, I target employment shares by occupational group at the 10, 50, and 90 wage percentiles to provide information on equilibrium sorting.

price polarization," meaning the relative prices of manual to routine skill and abstract to routine skills rise under task price polarization. I implement his estimator for task prices on model simulated cohorts over the same time period (1984/92-2007/09), substituting in my skill measures $x_C(0)$ and $x_M(0)$ and occupational groups in place of $(x_A(0), x_R(0), x_M(0))$ and his abstract, routine, and manual occupational groups.[90] The model simulated NLSY-like cohorts exhibit task price polarization despite wages arising from a markedly different framework and data construction (Table 9).

Autor and Dorn (2013) also consider a competitive Roy-style model and test its predictions about employment and wage changes in "routine" occupations.[91] They hypothesize that demand for routine tasks fell 1980 to 2005, causing areas with larger shares of routine occupations to experience drops in non-college worker wages and the share of routine employment. However, they find clerical occupations in commuting zones with higher routine employment shares experience a wage gain, weakening their results.[92] They hypothesize that self selection puts more productive workers in clerical jobs but cannot test this hypothesis due to a lack of data. The model can fill this gap and evaluate their hypothesis.

The model works in (y_C, y_M)-space instead of geographical space. Figure 22 shows routine intensity in (y_C, y_M)-space where lighter shading represents higher concentration.[93] I calculate changes in employment shares in "clerical" and "non-clerical" middle-skilled occupations in the model using its best fitting version (III).[94] Figure 22 shows moderate to high routine intensity for clerical jobs (red rectangle). From, 1980 to 2005, the employment share of "clerical" occupations fell slightly (-1%) while its average wage rose (+6%) in the model. This matches the pattern in the data even though the model does not target changes in this group over this period. Meanwhile, employment shares (-29%) and average wages (-3%) fell in the model's "non-clerical" medium-skilled occupational group. Workers selected into these occupation such that the average level of cognitive skills increases 4.7% in the "clerical" group versus 2.2% in the "non-clerical" group. Manual skills in each occupational group only changed slightly.

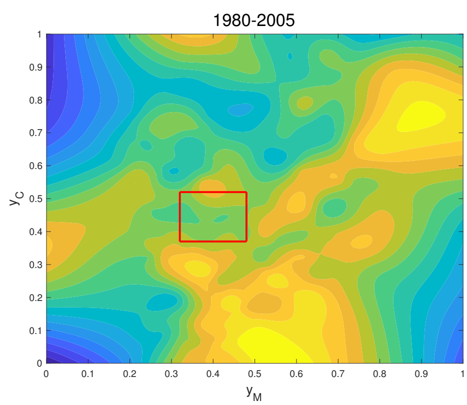


Figure 22: Routine Intensity in (y_C, y_M)-space

The hypothesis of Autor and Dorn (2013) says wages rise in clerical occupations due to displacement of the least skilled workers and most routine tasks within clerical occupations.[95] The increase in x_C supports

[90] Abstract, routine, and manual correspond to the high, medium, and low-skilled occuaptions in Acemoglu and Autor (2011). I selected the occupational groups to best align with these groups in a simple manner in Section 3.2.

[91] Routine job consists of repetatative, codifiable tasks, e.g. bank teller (Autor, Levy, and Murnane, 2003).

[92] The gain is not small. It is comparable to the percentage gain for low-skill service occupations and high-skill occupations.

[93] I calculate the Autor and Dorn (2013) measure of routine task intensity for occupations in 1979, map these occupations into (y_C, y_M)-space, and smooth over the contours.

[94] I define clerical occupations in (y_C, y_M)-space using the interquartile values of (y_C, y_M) estimated for clerical occupations in the data, i.e. $\{(y_C, y_M) : 0.37 < y_C < 0.52,\ 0.32 < y_M < 0.48\}$.

[95] The model provides a means to obtain measures of consistent within task changes in occupation that Autor and Dorn

the first part of their hypothesis. The second part of their hypothesis says the most routine-intensive tasks *within* clerical occupations become displaced. The change in the density $\mathcal{F}_t(\mathbf{y})$ negatively correlates (-0.25) with routine intensity within this occupational group. Thus, the skill demand distribution shifts away from areas holding initially higher routine intensity within the clerical group, supporting the second part of their hypothesis. This example and "task price polarization" show the model can produce results and test hypotheses from the literature, adding to the credibility of its own insights.

6 Drivers of Skill Demand

Now, I turn to evaluate the importance of various economic forces behind changes in skill demand. Various papers put forward strong candidates to explain why skill demand polarized, shifting away from middle-skilled occupations towards high and low-skilled occupations. Prominent explanations fall into the broad categories about technological progress, globalization, and consumer preferences. I map related variables into (y_C, y_M)-space and perform variance decompositions on $\mathcal{F}_t(\mathbf{y})$ to measure the relative importance of some of these explanations. The model delivers the whole (parameterized) distribution of skill demands from 1979 to 2010. This distribution provides the power and cross-sectional variation needed to identify the contributions of each variable considered.[96]

6.1 Data & Variables

The model casts occupation in terms of their task complexity, but prominent explanations also consider task content. Autor, Levy, and Murnane's (2003) routinization hypothesis claims forces like automation eliminated routine jobs. Goos, Manning, and Salomons (2014) use task content to extrapolate whether automation and offshoring account for job-polarizing skill demand changes.[97] I consider task content measures for offshoring vulnerability, routine-intensity, and interpersonal intensity estimated using O*NET and DOT via Autor and Dorn (2013).[98] Offshoring vulnerability measures the need for face-to-face contact and hence the ease of performing a task abroad. Figure 23 shows these jobs range from manually complex but cognitively simple tasks to cognitively complex and manually simple tasks. For example, insurance underwriters, $(y_C, y_M) = (0.67, 0.27)$ and machines operators, $(0.09, 0.60)$, fall into these categories. Routine-intensity measures the extent to which the job's tasks follow a set of codifiable rules and thereby susceptible to automation. Figure 23 shows these jobs consist of moderate to complex manual tasks of low cognitive complexity. Machine and telephone operators $(0.24, 0.4)$ serve as example of routine-intensive occupations. Interpersonal intensity measures the extent to which a job requires social skills like negotiation, persuasion, and emotional perception. Psychologists $(0.84, 0.30)$ serve as a good example of an interpersonal-intensive task. Visual evidence immediately suggests roles for offshoring and automation in explaining the decline in demand for medium-skilled occupations. Lighter areas in Figure 23 show high concentration of routine and offshorable tasks in 1979. These areas in (y_C, y_M)-space coincide with areas where skill demand declined

(2013) claim to need.

[96]The low frequency of the variables available makes spurious correlations likely without cross-sectional variation.

[97]They model technological change as linear time trends as I do in the model. Few datasets measuring realized automation and offshoring exists, which is why the literature uses task content as a proxy. One recent exception is Acemoglu and Restrepo (2017). They use a proprietary dataset to examine the role of robots.

[98]See Autor and Dorn (2013) for details on variable construction of offshoring vulnerability and routine intensity measures. Interpersonal intensity comes from O*NET measures for social perceptiveness, coordination, persuasion, negotiation, instruction, and service orientation.

the most (Figure 13).[99] These same areas lack intense use of interpersonal tasks while areas of increased demand use them intensively.

A variety of papers measure the impact of technology and trade via differences in technology adoption or trade exposure. These difference occur across industries, hence they exploit industry differences and variation in the industry mix across areas. Michaels, Natraj, and Van Reenen (2014) show large polarizing effects across industries due to accelerated information and communications technology (ICT) adoption and R&D using industry data across countries. Following Michaels, Natraj, and Van Reenen (2014), I use the flow of ICT expenditures as a share of value added to measure technological progress. In addition, I construct similar capital share variables for machinery, research and development (R&D), and transportation equipment. This data comes from the EU KLEMS Growth and Productivity Accounts Statistical Module. I also construct a measure of Chinese import penetration into manufacturing sub-sectors.[100] Autor, Dorn, and Hanson (2013) show large negative, local effects on manufacturing employment driven by rising import competition from China. Figure 24 shows the manufacturing industry concentrates in the area with the largest decline in skill demand. Manufacturing industry import and export data comes from Schott (2008). Data on domestic shipments comes from the NBER-CES Manufacturing Industry Database.[101] I aggregate these annual industry variables into 11 major sectors to create consistency in variables across time and datasets.[102] I then merge all of the above metrics into the CPS DOT dataset (Appendix B.3) based on these sectors to obtain variables over a (y_C, y_M) grid.[103] This approach mirrors Autor and Dorn (2013) and Autor, Dorn, and Hanson (2013). They use variation in local exposure to test predictions stemming changes in skill demand. I leverage the model and use variation in exposure across (y_C, y_M) to identify the impact of the factors mentioned on $\mathcal{F}_t(\mathbf{y})$.

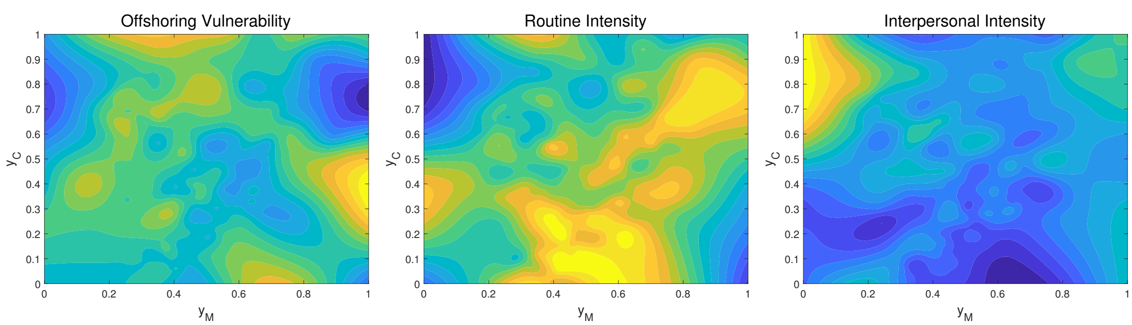


Figure 23: 1979 Task Content in (y_C, y_M)-space

To summarize, the factors I evaluate from the literature include ICT adoption, R&D, manufacturing import penetration, susceptibility to automation and vulnerability to offshoring. ICT, R&D, and vulnerability to automation constitute technological factors. Import penetration and offshoring risk serve as globalization

[99] I focus all the analysis here on estimated skill demands for the best fitting model (III). Additional results for other model versions are available upon request.

[100] Import penetration is the ratio of imports to net imports minus the total value of domestic shipments based on the definition of Lu and Ng (2013). Manufacturing sub-sectors are 1) food and tobacco, 2) textiles and appliances, 3) wood and furniture, 4) paper and printing, 5) chemicals and petroleum, 6) clay, stone, rubber and leather, 7) metals, 8) equipment, 9) transport, and 10) other products (e.g. toys).

[101] http://www.nber.org/nberces/. Accessed 28 July 2017.

[102] These sectors are 1) agriculture, forestry, fishing, and hunting, 2) mining, 3) construction, 4) manufacturing, 5) wholesale and retail trade, 6) transportation and utilities, 7) information and communications, 8) financial, professional and business services, 9) educational and health services, 10) leisure and hospitality, and 11) other services.

[103] I weight observations by the industry concentration within the respective CPS DOT occupation in a given year to obtain concentration variables over a (y_C, y_M) grid. I smooth these variables over the support of $\mathbf{y}$ which imputes values for jobs with similar task complexity but are unobserved in the data.

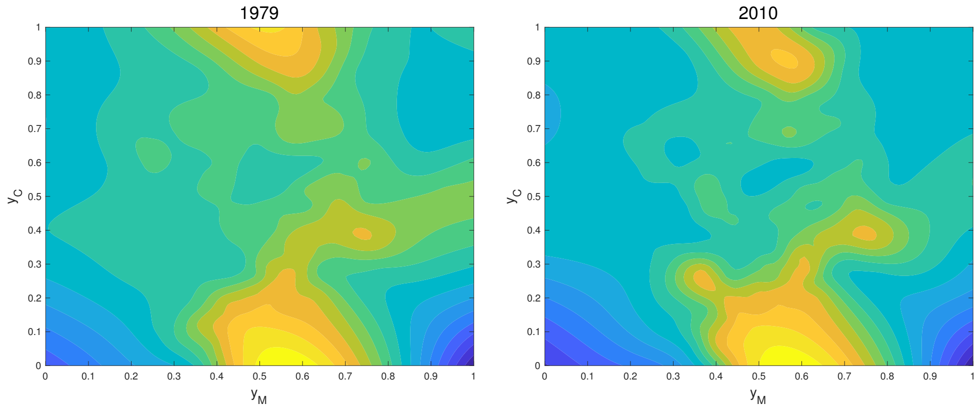


Figure 24: Manufacturing, Mining, and Construction Concentration in (y_C, y_M)-space

and trade related factors. None of these factors directly capture the effects of consumer preferences on skill demand. Autor and Dorn (2013) argue increased demand for low-skilled service occupations comes from the interaction of consumer preferences and technological change. Technological progress in goods production lowers their cost, but consumer prefer variety and thereby increase their demand for low-skilled, non-routine services. Similarly, firms performing highly complex tasks benefit from technological innovation and demand more of these services as they expand. Figure 25 shows the professional services industry provides jobs consisting of highly complex cognitive tasks and non-complex tasks. For example, receptionists perform relatively simple tasks, $(y_C, y_M) = (0.30, 0.22)$, and work mostly in the professional service industry. I capture this interaction by weighting industry level variables by their employment share within an occupation.

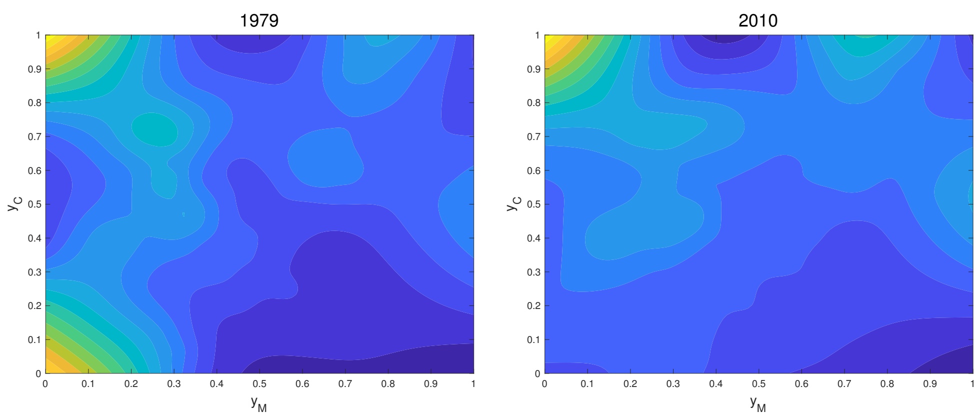


Figure 25: Financial, Professional and Business Service Concentration in (y_C, y_M)-space

6.2 Variance Decompositions

The literature tells us a myriad of factors significantly affected skill demand, contributing to polarization. However, the disparate nature of these studies makes evaluating their relative importance difficult. Here, the model proves useful. Its estimates of the distribution of skill demand in (y_C, y_M)-space provide a foundation to compare various factors once cast in this space. It implicitly provides variation within and across occupations over time. We can exploit the variation in changes across the support of (y_C, y_M) to measure which factors appear more important. I perform a simple linear variance decomposition on the change in $\mathcal{F}_t$ across (y_C, y_M) cells to measure each factors relative importance.[104] I focus on contributions rather than the significance of each factor, because the literature has established their significance.[105] But, it has not fully established their relative importance.

First, I examine in which industries changes in the distribution of skill demand took place given the visual evidence in Figures 24 and 25. Table 10 shows industry concentration in 1979 accounts for up to a half of the changes from 1979 to 2010.[106] In other words, industry (linear) trends alone account for half of the changes in the skill demand distribution. The manufacturing and construction industries account for much of industries' contribution as Figure 24 suggested. The rise in importance in information and professional services in the 1990s is consistent with rising ICT adoption as these industries experience the largest additions relative to total value added.

Next, I decompose changes in $\mathcal{F}_t(\mathbf{y})$ due to task content. Interpersonal intensity negatively correlates with both routine intensity (-0.66) and offshoring (-0.44) in 1979. Meanwhile, routine intensity and offshoring correlate weakly and positively (0.01). Table 11 shows the variance contributions of offshoring vulnerability, routine intensity, and interpersonal intensity to changes in the skill demand distribution. The fourth column shows interpersonal task content outweighs routine intensity and offshoring risk in importance. Ignoring interpersonal content overemphasizes the importance of routine content in accounting for changes in skill demand as the last column of Table 11 shows. They correlate strongly and negatively, but interpersonal intensity better accounts for $\Delta\mathcal{F}_t(\mathbf{y})$ by a factor of 7 to 1.[107] All else equal, higher interpersonal intensity at (y_C, y_M) correlates to increased skill demand (i.e. density of $\mathcal{F}$ rises). I interpret this correlation as demand increased for interpersonal skills over the three decades. Holding interpersonal intensity and offshoring risk fixed, demand decreased for routine skills but to a much lesser degree than it increased for interpersonal skills. These demand shifts appear to take place in the 1990s and 2000s, respectively. Automation may account for this fall in demand for routine skills as routine skills remain more susceptible to automation by their definition. Jobs with higher offshoring risk actually increase in demand, all else equal, especially in the 1990s.[108] These jobs include tasks which require high cognitive skills but little face-to-face contact like economists $(0.65, 0.25)$, accountants $(0.65, 0.23)$, and operations/systems analysts $(0.67, 0.34)$. Again, this increase outweighs the fall in routine skills in importance to account for $\Delta\mathcal{F}_t(\mathbf{y})$. Overall, the model's skill demand estimates do not reject Autor, Levy, and Murnane's routinization hypothesis. However, they emphasize asymmetry in the importance of the rise in demand for interpersonal skills versus the fall in demand for routine skills. The model's skill demands suggests industry trends in manufacturing and construction encompass most of the explanatory power of automation risk. Offshoring risk does not correlate to lower skill demand overall. In fact, demand rose for cognitively complex tasks at higher risk of offshoring, all else equal. Some of these jobs may have been offshored *to* the US, because the net flow of foreign direct investment (FDI) began to increase starting in 1990.

Finally, I turn to the industry level variables to provide insight how technology and trade account for $\Delta\mathcal{F}_t(\mathbf{y})$.[109] Table 12 presents the individual variance contribution of each factor and their joint contribution

[104] In Appendix C.2, I show that selection effect necessitate the use of the model primitive, $\mathcal{F}_t$, and not the equilibrium distribution of $\mathbf{y}$.

[105] The projection coefficients hold no meaningful interpretation with respect to changes in the density of $\mathcal{F}_t$.

[106] The total variance contribution displayed includes contributions due to correlations in industry concentration at (y_C, y_M) cells. I use a 100×100 grid for 10,000 cells.

[107] The partial R^2 of routine intensity is 1.6% compared to 11.8% for interpersonal intensity.

[108] The projection coefficient on offshoring vulnerability is positive for 1979 to 2010.

[109] I control for the initial industry shares in this decomposition, which is equivalent to including industry trends in the level regressions. I do this for the same reason as Autor, Dorn, and Hanson (2013). I want to use variation in industry level exposure (rather than industry trends) to identify the effects of changes in each factor.

to changes in the distribution of skill demand. The results suggest changes in machinery and transport adoption drove changes in demand in the 1980s. The increase in variance contribution for ICT confirms the finding of Michaels, Natraj, and Van Reenen (2014) for the 1990s. ICT drove changes in demand in the 1990s to a relatively large extent. The 2000s appear more mixed in what affects $\Delta\mathcal{F}_t(\mathbf{y})$. Overall, R&D and transport adoption appear to serve as the most important determinants of changes in skill demand over the three decades. The importance of manufacturing import competition from China diminishes over time as the manufacturing industry share falls.[110] ICT's impact occurred mainly in the 1990s. Industry trends, technological progress, and trade as measured by the variables shown explain up to 57% of the job-polarizing change in the distribution of skill demands from 1979 to 2010.[111]

I now provide a comprehensive interpretation of these results.[112] Continued productivity-enhancing (or labor-augmenting) industrialization in part drove the 1980s. Adoption of machinery made specific skills which use complex tasks more valuable and thereby increased their demand. At the same time, the manufacturing industry lowered demand for the manually complex tasks it performs (likely due to automation), forcing the least productive workers into low-skilled occupations.[113] Hence, we see job polarization but wage expansion across occupations and the distribution overall.[114] The accumulation of machinery also began to decelerate in the 1980s (Appendix Figure 54). The development of ICT in the 1990s created opportunities requiring high cognitive skill to leverage social skills like negotiation and persuasion. This key development led to occupational upgrading as demand shifted away from complex manual tasks towards complex cognitive tasks involving interpersonal skills. After the 1990s, it seems the impact of ICT development tapered. Automation susceptibility appears to have taken on some importance in the 2000s, but a lot of the changes in skill demand during this time remain unexplained.

7 Conclusion

This paper presents a quantitative model which reconciles changes in the occupational and wage structures. Reconciling these changes requires a framework which takes selection effects seriously. To this end, I employ a dynamic, multidimensional-skill search model. I use variation in micro data on wages, occupations, and task complexity to estimate model parameters and back out what skill demand shifts occurred. The model produces the observed patterns of expansion and contraction across occupations and the wage distribution over 1979 to 2010. It also reproduces some patterns observed in the reduce form literature on job and wage polarization. The model indicates selection based on heterogenous specific human capital plays an important role in accounting for the observed allocation of workers to jobs.

I then take the estimated shifts in the distribution of skill demands and use them to evaluate explanations for these changes. I find industry trends, technological progress, and trade as measured by the variables shown explain up to 57% of the polarization in the distribution of skill demand over 1979 to 2010. Looking closer, the adoption of machinery, transport equipment and R&D appear to hold some importance throughout the three decades. However, ICT adoption took on a strong role in the 1990s and spurred demand for interpersonal and social skills. The results suggest this "ICT Revolution" changed the occupational and

[110] Autor, Dorn, and Hanson (2013) instrument this variable, but obtain similar results with OLS and 2SLS.
[111] The remaining 43% and lack of explanatory power in the 1980s and 2000s prompt questions beyond the scope of this paper.
[112] Of course, this interpretation does not rule out others.
[113] The decline in manufacturing employment share in the 1980s onward look to be part of a long-run trend. See <https://fred.stlouisfed.org/series/USAPEFANA>.
[114] Workers with few skills (often younger) tend to make more gains through experience. During the 1980s, life cycle wage profiles flattened and began to become stepper again more recently Manovskii and Kambourov (2005). This occurance likely relates to occupational wage polarization.

Table 10: Initial Industry Concentration Variance Decomposition on $\Delta\mathcal{F}(\mathbf{y})$

	1979-1989	1989-2000	2000-2010	1979-2010
Agriculture, Forestry, Fishing, & Hunting	0.033	0.006	0.045	0.001
Mining	0.001	0.001	0.003	0.000
Construction	0.063	0.026	0.010	0.074
Manufacturing	0.020	0.098	0.056	0.125
Wholesale & Retail Trade	0.000	0.024	0.001	0.011
Transportation & Utilities	0.000	0.045	0.029	0.010
Information Services	0.010	0.048	0.001	0.034
Financial, Professional, & Business Services	0.003	0.041	0.007	0.068
Education and Health Services	0.004	0.018	0.003	0.012
Leisure & Hospitality	0.020	0.001	0.011	0.000
Other Services	0.002	0.008	0.000	0.008
Total Variance Contribution (R^2)	13.8%	37.4%	16.0%	47.1%

Table 11: Initial Task Content Variance Decomposition on $\Delta\mathcal{F}(\mathbf{y})$

	1980s	1990s	2000s	1979-2010	1979-2010
Offshoring Vulnerability	0.040	0.164	0.009	0.128	0.021
Routine Intensity	0.004	0.001	0.090	0.023	0.226
Interpersonal Intensity	0.025	0.400	0.002	0.239	–
Total Variance Contribution (R^2)	2.8%	30.2%	7.4%	33.5%	24.6%

Table 12: Capital Input and Imports Variance Decomposition on $\Delta\mathcal{F}(\mathbf{y})$

	1980s	1990s	2000s	1979-2010
Individual Contributions				
Δ Chinese Manufacturing Import Penetration	0.082	0.009	0.005	0.003
Δ Capital Formation				
ICT	0.006	0.176	0.031	0.006
Machinery	0.087	0.006	0.062	0.039
R&D	0.007	0.009	0.021	0.094
Transportation Equipment	0.162	0.064	0.058	0.104
Joint Contribution (Partial R^2)	7.5%	18.6%	4.9%	16.3%
Total Variance Contribution (with industry mix)	20.3%	43.2%	20.1%	56.9%

wage structure far more than the decline in demand for routine skills. Shifts in routine skills do not appear quantitatively important outside of the long-run decline in manufacturing and construction employment. Still, the variables used fail to account for about 40% of the changes in the distribution of skill demand, leaving several questions. Is what is left truly unexplained noise at the occupational level or estimation error in $\Delta\mathcal{F}_t(\mathbf{y})$ or noise due to aggregating to the sector level? Are there technological changes yet to be widely considered (e.g. robots, better measures of automation) which account for much of what is left? All of these questions remain important for future work of this nature.

The model points to several avenues of potential research. First, the model takes a first step at quantitatively introducing frictions and dynamic selection issues neglected in the job/wage polarization literature. In doing so, it takes meeting rates as given to deal with the inherently non-stationary nature of the transition path. As a result, the model remains silent on the role of search related general equilibrium feedback. In this dimension, further developments in directed search may prove fruitful. For example, Menzio and Shi (2010) present a block recursive directed search model which makes meeting rates independent of the endogenous distribution of worker types. However, the model has no firm heterogeneity. Thus, it also has no meaningful notion of skill mismatch or imperfect sorting, which arguably remains an important feature of the data. Second, the model views labor market structural change as the result of a gradual process. Hershbein and Kahn (2016) provide evidence which says short downturns amplify the negative impact of these changes on (cognitive) routine skill demand. The exogenous nature of skill demands and free entry here means the model cannot speak to the timing of structural adjustment on the firm side. Understanding this requires progress in understanding how firms determine their multidimensional skill demands.

References

Acemoglu, D., and D. Autor (2011): *Skills, Tasks and Technologies: Implications for Employment and Earnings*vol. 4 of *Handbook of Labor Economics*, chap. 12, pp. 1043–1171. Elsevier.

Acemoglu, D., D. Autor, D. Dorn, G. H. Hanson, and B. Price (2016): "Import Competition and the Great US Employment Sag of the 2000s," *Journal of Labor Economics*, 34(S1), 141–198.

Acemoglu, D., and P. Restrepo (2017): "Robots and Jobs: Evidence from US Labor Markets," Boston University - Department of Economics - Working Papers Series dp-297, Boston University - Department of Economics.

Altonji, J. G., P. Bharadwaj, and F. Lange (2012): "Changes in the Characteristics of American Youth: Implications for Adult Outcomes," *Journal of Labor Economics*, 30(4), 783–828.

Autor, D. H. (2015): "Why Are There Still So Many Jobs? The History and Future of Workplace Automation," *Journal of Economic Perspectives*, 29(3), 3–30.

Autor, D. H., and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5), 1553–1597.

Autor, D. H., D. Dorn, and G. H. Hanson (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103(6), 2121–2168.

Autor, D. H., and M. J. Handel (2013): "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages," *Journal of Labor Economics*, 31(S1), 59–96.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?," *The Quarterly Journal of Economics*, 113(4), 1169–1213.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content Of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118(4), 1279–1333.

Bartel, A., C. Ichniowski, and K. Shaw (2007): "How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills," *The Quarterly Journal of Economics*, 122(4), 1721–1758.

Boehm, M. (2017): "The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change," mimeo, University of Bonn.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002): "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *The Quarterly Journal of Economics*, 117(1), 339–376.

Bureau of Labor Statistics (2016): "Characteristics of minimum wage workers, 2015," Discussion Paper 1061, U.S. Bureau of Labor Statistics, Washington, DC, An optional note.

——— (2017): "Occupation Requirements Survey," `https://www.bls.gov/ors/`.

Bureau of Labor Statistics, U.S. Department of Labor (2014a): *National Longitudinal Survey of Youth 1979 cohort, 1979-2012 (rounds 1-25),* Produced and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

——— (2014b): *National Longitudinal Survey of Youth 1997 cohort, 1997-2013 (rounds 1-16),* Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, 74(2), 323–364.

Center for Economic and Policy Research (2017): *CPS ORG Uniform Extracts, Version 2.2.1,* Washington, DC.

Chari, V. V., and H. Hopenhayn (1991): "Vintage Human Capital, Growth, and the Diffusion of New Technology," *Journal of Political Economy*, 99(6), 1142–1165.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2016): "Disappearing Routine Jobs: Who, How, and Why?," NBER Working Papers 22918, National Bureau of Economic Research, Inc.

Dorn, D. (2009): "Essays on Inequality, Spatial Interaction, and the Demand for Skills," Dissertation, University of St. Gallen.

England, P., and B. Kilbourne (1980): *Occupational Measures from the Dictionary of Occupational Titles for 1980 Census Detailed Occupations,* Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013-06-20.

Eriksson, T., and N. Kristensen (2014): "Wages or Fringes? Some Evidence on Trade-Offs and Sorting," *Journal of Labor Economics*, 32(4), 899–928.

Firpo, S., N. M. Fortin, and T. Lemieux (2011): "Occupational Tasks and Changes in the Wage Structure," IZA Discussion Papers 5542, Institute for the Study of Labor (IZA).

Gibbons, R., and M. Waldman (2004): "Task-Specific Human Capital," *American Economic Review*, 94(2), 203–207.

Goos, M., A. Manning, and A. Salomons (2014): "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, 104(8), 2509–2526.

Hagedorn, M., T. H. Law, and I. Manovskii (2017): "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85(1), 29–65.

Hawkins, W. B., and D. Acemoglu (2014): "Search with multi-worker firms," *Theoretical Economics*, 9(3).

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, 13(1), 15–51.

Heckman, J., L. Lochner, and C. Taber (1998): "Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58.

Hershbein, B., and L. B. Kahn (2016): "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings," NBER Working Papers 22762, National Bureau of Economic Research, Inc.

IPUMS-CPS, University of Minnesota (2016): *Current Population Survey Data,* `http://www.ipums.org`. Accessed: 2016-08-23.

Jaimovich, N., H. Siu, and G. M. Cortes (2017): "The End of Men and Rise of Women in the High-Skilled Labor Market," Discussion paper.

Jones, M. (2009): "Kumaraswamy's distribution: A beta-type distribution with some tractability advantages," *Statistical Methodology*, 6(1), 70–81.

Kambourov, G., and I. Manovskii (2002): "Occupational Specificity of Human Capital," mimeo, University of Pennsylvania.

Kantenga, K., and T.-H. Law (2017): "Sorting and Wage Inequality," mimeo, The University of Pennsylvania.

Katz, L. F., and D. H. Autor (1999): "Changes in Wage Structure and Earnings Inequality," in *Handbook of Labor Economics*, ed. by O. Ashtenfelter, and D. Card, vol. 3A, pp. 1463–1555. Amsterdam: North Holland.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

Kredler, M. (2014): "Experience vs. obsolescence: A vintage-human-capital model," *Journal of Economic Theory*, 150(C), 709–739.

Lefter, A., and B. M. Sand (2011): "Job Polarization in the U.S.: A Reassessment of the Evidence from the 1980s and 1990s," Economics Working Paper Series 1103, University of St. Gallen, School of Economics and Political Science.

Lemieux, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 96(3), 461–498.

Lindenlaub, I. (2017): "Sorting Multidimensional Types: Theory and Application," *Review of Economic Studies*, 84(2), 718–789.

Lise, J., and F. Postel-Vinay (2016): "Multidimensional Skills, Sorting, and Human Capital Accumulation," Revise and resubmit.

Lu, Y., and T. Ng (2013): "Import Competition and Skill Content in U.S. Manufacturing Industries," *The Review of Economics and Statistics*, 95(4), 1404–1417.

Manovskii, I., and G. Kambourov (2005): "Accounting for the Changing Life-Cycle Profile of Earnings," Discussion paper.

Menzio, G., and S. Shi (2010): "Directed Search on the Job, Heterogeneity, and Aggregate Fluctuations," *American Economic Review*, 100(2), 327–332.

——— (2011): "Efficient Search on the Job and the Business Cycle," *Journal of Political Economy*, 119(3), 468–510.

Menzio, G., I. Telyukova, and L. Visschers (2016): "Directed Search over the Life Cycle," *Review of Economic Dynamics*, 19, 38–62.

Michaels, G., A. Natraj, and J. Van Reenen (2014): "Has ICT polarized skill demand? Evidence from eleven countries over 25 years," LSE Research Online Documents on Economics 46830, London School of Economics and Political Science, LSE Library.

Miller, A., D. Treiman, P. Cain, and P. Roos (1980): *Work, Jobs, and Occupations: A Critical Review of the Dictionary of Occupational Titles*. National Academy Press, Washington, DC.

Mishel, L., J. Bivens, E. Gould, and H. Shierholz (2012): *The State of Working America*. Cornell University Press, Ithaca, New York, 12 edn.

Mishel, L., J. Schmitt, and H. Shierholz (2013): "Assessing the Job Polarization Explanation of Growing Wage Inequality," Epi-cepr working paper, Economic Policy Institute.

Mortensen, D. T., and C. A. Pissarides (1999): "New developments in models of search in the labor market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 39, pp. 2567–2627. Elsevier.

Moscarini, G., and K. Thomsson (2006): "Occupational and Job Mobility in the US," Working Papers 19, Yale University, Department of Economics.

National Bureau of Economic Research (2016): "CPS Merged Outgoing Rotation Groups," `http://www.nber.org/data/morg.html` Accessed: 2016-08-23.

National Crosswalk Service Center (2016): "O*NET Resource Center," `http://www.xwalkcenter.org` Accessed: 2016-08-23.

O*NET (2016): "O*NET Resource Center," `https://www.onetcenter.org` Accessed: 2016-08-23.

Postel-Vinay, F., and G. Moscarini (2009): "Non-Stationary Search Equilibrium," Discussion paper.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

Robin, J.-M. (2009): "Labour Market Dynamics with Sequential Auctions and Heterogeneous Workers," mimeo, University College London.

Roy, A. D. (1951): "Some Thoughts On The Distribution Of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

Sanders, C. (2012): "Skill Uncertainty, Skill Accumulation, and Occupational Choice," Discussion paper.

——— (2016): "Skill Uncertainty, Skill Accumulation, and Occupational Choice," Discussion paper.

Sanders, C., and C. Taber (2012): "Life-Cycle Wage Growth and Heterogeneous Human Capital," *Annual Review of Economics*, 4(1), 399–425.

Schmitt, J. (2003): "Creating a consistent hourly wage series from the Current Population Survey's Outgoing Rotation Group, 1979-2002," mimeo, Center for Economic and Policy Research.

Schott, P. K. (2008): "The relative sophistication of Chinese exports," *Economic Policy*, 23, 5–49.

Segall, D. O. (1997): "Equating the CAT-ASVAB," in *Computerized Adaptive Testing: From Inquiry to Operation*, ed. by W. A. Sands, B. K. Waters, and J. R. McBride, chap. 19, pp. 181–198. American Psychological Association, Washington, DC, 1 edn.

Sherk, J. (2013): "Productivity and Compensation: Growing Together," Discussion Paper 2825, The Heritage Foundation, Washington, DC, An optional note.

Shimer, R. (2012): "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics*, 15(2), 127–148.

Slonimczyk, F. (2013): "Earnings inequality and skill mismatch in the U.S.: 1973âĂŞ2002," *The Journal of Economic Inequality*, 11(2), 163âĂŞ194.

U.S. Department of Labor (1991): *The Revised Handbook for Analyzing Jobs*. JIST Works, Indianapolis, IN.

U.S. Department of Labor, U.S. Employment Service, and the North Carolina Occupational Analysis Field Center (1991): *DICTIONARY OF OCCUPATIONAL TITLES (DOT): REVISED FOURTH EDITION,*

Washington, DC: United States Department of Labor, United States Employment Service, and Raleigh, NC: North Carolina Occupational Analysis Field Center [producer], 1991. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1994.

Valletta, R. G. (2016): "Recent Flattening in the Higher Education Wage Premium: Polarization, Skill Downgrading, or Both?," IZA Discussion Papers 10194, Institute for the Study of Labor (IZA).

Willis, R. J., and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87(5), 7–36.

Yamaguchi, S. (2012): "Tasks and Heterogeneous Human Capital," *Journal of Labor Economics*, 30(1), 1 – 53.

Appendix (For Online Publication)

A Model Appendix

A.1 Occupational Wages, Employment Shares, and Wages

Typically, workers in the low, medium and high wage occupational groups populate the 10th, 50th, and 90th wage percentiles, respectively (Appendix Figure 49). Wages overall reflect wage changes between these occupation groups. Thereby, wage polarization between occupational groups may induce wage polarization overall. However, neither job polarization nor occupational wage polarization serve as necessary nor sufficient conditions to generate wage polarization. Workers switching ranks in the wage distribution may undo wage polarization. To illustrate, suppose there exists 3 workers in low, medium and high-paying jobs. The high-earning worker earns 10 units, the medium worker earns 5 units, and the low worker earns 4 units. Now suppose the low worker gains 3 units, the high worker gains 10, and the medium worker gains none. The low worker earns 75% more and the high worker earns 50% more. Clearly, occupation wage polarization occurs. However, the change in wages for the 1st ranked worker is only 20% (4→5), the change for the 2nd ranked worker is 40% (5→7), and the change for the highest ranked worker remains 50%. Wages overall do not polarize even though occupational wages polarize, because the low worker overtakes the medium worker. If the low worker gains any amount within the interval (0, 2), then wages polarize. Such complications necessitate using the model to isolate what generates wage polarization in the 1990s.

Clearly, occupational wage polarization is not sufficient to ensure wage polarization. Furthermore, it is not even necessary. Wages may polarize mainly due to changes in the occupational distribution. Suppose the medium occupation expands to populate the 10th percentile and its wage increases slightly. Meanwhile, the low occupation shrinks and its wages stagnate. Call this phenomenon "occupational upgrading." The increased presence of medium earners compresses wages in the bottom half of the wage distribution. The higher wage now at the 10th percentile disproportionately increases the wage growth at lower percentiles while only increasing wage growth slightly at the 50th percentile. If wages in the high occupational group grow enough at the 90th percentile, then wage polarization occurs but neither occupational wage polarization nor job polarization occur. Again, these complications necessitate using the model to isolate what generates the patterns of wage expansion and polarization we observe.

A.2 Bargaining Protocol

The bargaining protocol described in Section 2.1.2 serves several purposes. First, it is empirically relevant. Cahuc, Postel-Vinay, and Robin (2006) present evidence that this intra-employer competition or "job ladder" effect matters for wage determination. Second, this effect may also be important to explain changes across the wage and occupational wage distribution. For example, the average wage in a low-skilled occupation may rise due to a disproportionate number of workers in that occupation climbing the job ladder. We may attribute a wage increase due to wage dynamics to demand shifts if we ignore this effect. Third, this protocol delivers a value to the worker beyond their outside option, which enables the model to generate more realistic wage dynamics and wage levels. In Lise and Postel-Vinay (2016), even highly skilled workers can receive low and perhaps negative wages when the surplus is sufficiently large. In such a case, the high-skilled worker stands to gain significantly upon a job-to-job move due to the intra-employer competition. Consequently, the employer has a strong incentive to backload wage payments as much value to the high-skilled worker will be delivered upon a job-to-job transition. In fact, Lise and Postel-Vinay (2016) drop wages out of unemployment in their estimation due to the strength of this mechanism.[115] I give the worker some explicit bargaining power ($\lambda > 0$) to dampen this effect.

This bargaining protocol also which distinguishes this model from Lise and Postel-Vinay (2016) who use the bargaining protocol of Postel-Vinay and Robin (2002). Future gains beyond $S(\mathbf{x}, \mathbf{y})$ accrue to the new employer Lise and Postel-Vinay (2016), because λ is zero. This assumption results in the worker and the employer not taking the gains from a future job-to-job move into account when determining whether to form the match. Here, expected gains from future moves not only affect wages today as in Lise and Postel-Vinay (2016), but they also affect the job selection decision of the worker.[116] This occurs because the worker's expectation over future gains from a move affects the continuation value of a match when λ is not zero. In this way, workers care about their potential career path when accepting and declining job offers. Furthermore, employers care about the risk of a worker being poached when forming a match. Whereas in Lise and Postel-Vinay (2016), the match value does not take future moves into account, because the gains from a job-to-job move accrue entirely to the new employer.[117]

A.3 Surplus and Wages with No Offer Writing Cost

In this appendix, I derive the surplus and wage function in the case where a meeting that fails to deliver a job-to-job transition may still bid up the wage at the current employer. In this case, the renegotiated share of the surplus (σ') is

$$\sigma' = \sigma(\mathbf{x}, \mathbf{y}, \mathbf{y}') = \lambda + (1-\lambda)\frac{S(\mathbf{x}, \mathbf{y}')}{S(\mathbf{x}, \mathbf{y})} \in (0, 1]. \tag{19}$$

and the value to the employed worker is

$$\begin{aligned} W_t(\mathbf{x}, \mathbf{y}, \sigma) = \quad & w_t(\mathbf{x}, \mathbf{y}, \sigma) - c(\mathbf{x}, \mathbf{y}) + \beta_a \mathbb{E}_t U_{t+1}(\mathbf{x}') + \beta_a(1-\delta)(1-\mathbb{M}_{e,t})\sigma \mathbb{E}_t \tilde{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \\ & \beta_a(1-\delta)\mathbb{M}_{e,t} \times \\ & \mathbb{E}_t \int_{\mathcal{Y}} \max\{\lambda S_{t+1}(\mathbf{x}', \mathbf{y}) + (1-\lambda)\widehat{S}_{t+1}(\mathbf{x}', \mathbf{y}'), \lambda S_{t+1}(\mathbf{x}', \mathbf{y}') + (1-\lambda)\widehat{S}_{t+1}(\mathbf{x}', \mathbf{y})]\} \mathrm{d}\mathcal{F}_t(\mathbf{y}') \end{aligned} \tag{20}$$

where

$$\begin{aligned} \widehat{S}_{t+1}(\mathbf{x}, \mathbf{y}) =& \max\{S_{t+1}(\mathbf{x}, \mathbf{y}), 0\} \\ \tilde{S}_{t+1}(\mathbf{x}', \mathbf{y}) =& (1-\omega)\widehat{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \omega \int_{\mathcal{Y}} \max\{S_{t+1}(\mathbf{x}', \mathbf{y}'), 0\} \mathrm{d}\mathcal{F}_t(\mathbf{y}') \end{aligned}$$

[115] See Footnote 25 in Lise and Postel-Vinay (2016).

[116] Risk neutrality (i.e. linear preferences) makes the total surplus independent of the time profile of wage payments in Lise and Postel-Vinay (2016). Workers accept and reject offers based on the total surplus which does not depend on expectations over future gains from offers on-the-job.

[117] In Lise and Postel-Vinay (2016), workers and employers do care about how workers' skills evolve as a result of forming the match as they do here. In this way, workers care about their potential skill evolution when selecting a job. However, the path of future skill requirements does not affect the value of match and thus does not affect job selection.

subject to (1). We can now derive the following surplus and wages as in the main section. The value of producing now solves

$$
\begin{aligned}
P_t(\mathbf{x},\mathbf{y},\sigma) \quad &= f_t(\mathbf{x},\mathbf{y}) - w_t(\mathbf{x},\mathbf{y},\sigma) + \beta_a(1-\delta)(1-\mathbb{M}_{e,t})(1-\sigma)\mathbb{E}_t\tilde{S}_{t+1}(\mathbf{x}',\mathbf{y}) + \\
&\beta(1-\delta)\mathbb{M}_{e,t}(1-\lambda)\mathbb{E}_t\int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}',\mathbf{y}) - \widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}')\}\mathrm{d}\mathcal{F}_t(\mathbf{y}'). \qquad (21)
\end{aligned}
$$

We can show that the surplus which follows is

$$
\begin{aligned}
S_t(\mathbf{x},\mathbf{y}) = \quad & f_t(\mathbf{x},\mathbf{y}) - c(\mathbf{x},\mathbf{y}) - b(\mathbf{x}) + \beta_a(1-\delta)\mathbb{E}_t\Big[-\lambda\mathbb{M}_{u,t}\int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\ \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})\ + \\
& \big(1-\mathbb{M}_{e,t}\big)\cdot\max\{0,\tilde{S}_{t+1}(\mathbf{x}',\mathbf{y})\} + \\
& \mathbb{M}_{e,t}\cdot\max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\} + \\
& \mathbb{M}_{e,t}\cdot(1-\rho(\mathbf{x},\mathbf{y}))\cdot\Big[\lambda\big(\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}) - \widehat{S}_{t+1}(\mathbf{x}',\mathbf{y})\big)\Big]\Big], \qquad (22)
\end{aligned}
$$

$$
\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}) = \frac{\int_{\mathcal{Y}} \mathbb{1}\{\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) < S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\cdot S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\ \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}{\int_{\mathcal{Y}} \mathbb{1}\{\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) < S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}.
$$

(22) is identical to (12). Intuitively, the surplus should not change, because this modification changes the *split* of the surplus but not the surplus itself. If employers (who draw $\mathbf{y}'$ such that they cannot poach the employer) engage in Bertrand competition anyway, then they merely bid up the share of the surplus the worker's receives in the current match. These employers do not change the total value of the current match. However, wages depend on how the worker and employer split the surplus. Hence, the wage function changes to the following.

$$
\begin{aligned}
w_t(\mathbf{x},\mathbf{y},\sigma) = \quad & \sigma f_t(\mathbf{x},\mathbf{y}) + (1-\sigma)c(\mathbf{x},\mathbf{y}) + (1-\sigma)b(\mathbf{x}) + (1-\sigma)\beta_a(1-\delta)\times \\
& \mathbb{E}_t\Big[\lambda\mathbb{M}_{u,t}\int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\}\ \mathrm{d}\mathcal{F}_t(\mathbf{y}) - \mathbb{M}_{e,t}\cdot\max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\} - \\
& \mathbb{M}_{e,t}\cdot(1-\rho(\mathbf{x},\mathbf{y}))\cdot\lambda\cdot\Big(\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}') - \widehat{S}_{t+1}(\mathbf{x}',\mathbf{y})\Big)\Big] + \\
& \beta_a(1-\delta)\mathbb{M}_{e,t}(1-\lambda)\mathbb{E}_t\Big[\rho(\mathbf{x},\mathbf{y})\cdot\Big(S_{t+1}(\mathbf{x}',\mathbf{y}') - \underline{S}_{t+1}(\mathbf{x}',\mathbf{y}')\Big)\Big]. \qquad (23)
\end{aligned}
$$

where

$$
\underline{S}_{t+1}(\mathbf{x}',\mathbf{y}) = \frac{\int_{\mathcal{Y}} \mathbb{1}\{S_{t+1}(\mathbf{x}',\mathbf{y}) \geq \widehat{S}_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\cdot S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\ \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}{\int_{\mathcal{Y}} \mathbb{1}\{S_{t+1}(\mathbf{x}',\mathbf{y}) \geq \widehat{S}_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}.
$$

We can also rewrite (13) as

$$
\begin{aligned}
w_t(\mathbf{x},\mathbf{y},\sigma) = \quad & \sigma f_t(\mathbf{x},\mathbf{y}) + (1-\sigma)c(\mathbf{x},\mathbf{y}) + (1-\sigma)b(\mathbf{x}) + (1-\sigma)\beta_a(1-\delta)\times \\
& \mathbb{E}_t\Big[\lambda\mathbb{M}_{u,t}\int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\}\ \mathrm{d}\mathcal{F}_t(\mathbf{y}) - \mathbb{M}_{e,t}\cdot\max\{S_{t+1}(\mathbf{x}',\mathbf{y}),0\} - \\
& \mathbb{M}_{e,t}\cdot(1-\rho(\mathbf{x},\mathbf{y}))\cdot\lambda\cdot\Big(\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}') - \widehat{S}_{t+1}(\mathbf{x}',\mathbf{y})\Big)\Big] + \\
& \beta_a(1-\delta)\mathbb{M}_{e,t}(1-\sigma)\mathbb{E}_t\big[\rho(\mathbf{x},\mathbf{y})\cdot\max\{S_{t+1}(\mathbf{x}',\mathbf{y}),0\}\big]. \qquad (24)
\end{aligned}
$$

(24) is almost identical to (13) but $(1-\lambda)\mathbb{M}_{e,t}\rho(S-\underline{S})$ replaces $(1-\sigma)\mathbb{M}_{e,t}\rho S$ in the continuation value. The difference in these terms comes from the difference in the bargaining protocols with and without offer writing costs. Without offer writings costs, the potential outside offers which will not steal the worker still affect the value of the surplus delivered to the worker in the current match. These potential offers affect the worker's expected share of the surplus tomorrow and hence affect the value of the current match. Consequently, the wage adjusts downward when the expected value for a bidding up offer ($\underline{S}$) increases to deliver the surplus split σ today. In essence, the worker takes lower wages today with a greater expectation that the wage will be bid up on the job. With offer writing costs, no employer who draws skill requirements with lower surplus than the current match bids for the worker and hence the worker's value in the current match remains the same. As mentioned in section 2.2, this restriction prevents bidding up of wages on-the-job in order to restrict attention to human capital evolution over job shopping in the model. Both mechanisms can produce wage growth over a job's tenure. However quantitatively, the job shopping mechanism tends to generate large, counterfactual wage jumps on-the-job compared to more gradual wage changes due to human capital evolution. It also generates counterfactually low and negative wages due to promises of the wages being bid up over the job tenure. Hence, I assume the offering writing cost to preclude these counterfactuals.

A.4 Nash Bargaining Protocol

In this section, all workers with power $\lambda \in [0,1]$ bargain with employers à la Nash. Again, I assume the share of the surplus stays constant until an on-the-job meeting triggers renegotiation. I also assume unemployed workers accept job offers when indifferent. Hence, workers take a share of the surplus equal to λ. Again, a job-to-job transition only occurs when the surplus for the poaching employer exceeds that of the current employer. The unemployed worker's value function $U_t(\mathbf{x})$ imposing the bargaining protocol does not change as all unemployed workers Nash bargain in the benchmark model. The employed worker's value function $W_t(\mathbf{x},\mathbf{y},\lambda)$, imposing the bargaining protocol, solves

$$\begin{aligned} W_t(\mathbf{x},\mathbf{y}) \quad &= w_t(\mathbf{x},\mathbf{y}) - c(\mathbf{x},\mathbf{y}) + \beta_a \mathbb{E}_t U_{t+1}(\mathbf{x}') + \beta_a(1-\delta)(1-\mathbb{M}_{e,t})\lambda \mathbb{E}_t \tilde{S}_{t+1}(\mathbf{x}',\mathbf{y}) + \\ &\qquad \beta_a(1-\delta)\mathbb{M}_{e,t} \times \\ &\qquad \lambda \mathbb{E}_t \int_{\mathcal{Y}} \max\{\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}), S_{t+1}(\mathbf{x}',\mathbf{y}')\} \mathrm{d}\mathcal{F}_t(\mathbf{y}'), \end{aligned} \tag{25}$$

where

$$\begin{aligned} \widehat{S}_{t+1}(\mathbf{x},\mathbf{y}) &= \max\{S_{t+1}(\mathbf{x},\mathbf{y}), 0\} \\ \tilde{S}_{t+1}(\mathbf{x}',\mathbf{y}) &= (1-\omega)\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) + \omega \int_{\mathcal{Y}} \max\{S_{t+1}(\mathbf{x}',\mathbf{y}'), 0\} \mathrm{d}\mathcal{F}_t(\mathbf{y}'). \end{aligned}$$

The value of a vacancy V_t solves

$$\begin{aligned} V_t \quad &= -\tau_t + (1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{u,t}\lambda \mathbb{E}_t \int_{\mathcal{Y}} \int_{\mathcal{X}|u} \beta_a \max\{0, S_{t+1}(\mathbf{x},\mathbf{y})\} \mathrm{d}\mathcal{F}_t(\mathbf{y}) \mathrm{d}\mathcal{W}_t(\mathbf{x}|u) + \\ &\qquad (1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{e,t}(1-\lambda) \times \\ &\qquad \mathbb{E}_t \int_{\mathcal{Y}} \int_{\mathcal{X}|e} \beta_a \max\{0, S_{t+1}(\mathbf{x},\mathbf{y})\} \mathrm{d}\mathcal{F}_t(\mathbf{y}) \mathrm{d}\mathcal{W}_t(\mathbf{x}|e) \end{aligned} \tag{26}$$

where $\mathcal{W}_t(\mathbf{x}|u)$ and $\mathcal{W}_t(\mathbf{x}|e)$ are the distributions of unemployed and employed workers at time t, respectively. (26) differs from (7) in that employer do not need to compute expectations over all matches, because the bargaining power stays the same for all workers. I assume free entry of employers which drives the value of vacancy to zero so that

$$\begin{aligned}\tau_t = \quad & (1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{u,t}\lambda\mathbb{E}_t \int_{\mathcal{Y}}\int_{\mathcal{X}|u} \beta_a \max\{0, S_{t+1}(\mathbf{x},\mathbf{y})\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|u) + \\ & (1-\delta)\mathbb{M}_{v,t}\mathbb{C}_{e,t}(1-\lambda) \times \\ & \mathbb{E}_t \int_{\mathcal{Y}}\int_{\mathcal{X}|e} \beta_a \max\{0, S_{t+1}(\mathbf{x},\mathbf{y})\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|e).\end{aligned} \tag{27}$$

The value of producing solves

$$\begin{aligned}P_t(\mathbf{x},\mathbf{y}) \quad &= f_t(\mathbf{x},\mathbf{y}) - w_t(\mathbf{x},\mathbf{y}) + \beta_a(1-\delta)(1-\mathbb{M}_{e,t})(1-\lambda)\mathbb{E}_t\tilde{S}_{t+1}(\mathbf{x}',\mathbf{y}) + \\ & \beta_a(1-\delta)\mathbb{M}_{e,t}(1-\lambda)\mathbb{E}_t\big[\max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\}\cdot\rho(\mathbf{x},\mathbf{y})\big]\end{aligned} \tag{28}$$

where

$$\rho(\mathbf{x},\mathbf{y}) = \int_{\mathcal{Y}} \mathbb{1}\{S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}}) < S_{t+1}(\mathbf{x}',\mathbf{y})\}\mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}}).$$

$\rho(\mathbf{x},\mathbf{y})$ is the probability the worker at $\mathbf{y}$ does not draw an employer with higher surplus. $\mathbb{1}\{\cdot\}$ denotes the indicator function. We can now derive the surplus function using (5), (25), (28), and the free entry assumption which implies that V_t equals zero. For non-retiring workers, the surplus is

$$\begin{aligned}S_t(\mathbf{x},\mathbf{y}) = \quad & f_t(\mathbf{x},\mathbf{y}) - c(\mathbf{x},\mathbf{y}) - b(\mathbf{x}) + \beta_a(1-\delta)\mathbb{E}_t\Bigg[-\lambda\mathbb{M}_{u,t}\int_{\mathcal{Y}}\max\{0, S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\ \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})\ + \\ & \big(1-\mathbb{M}_{e,t}\big)\tilde{S}_{t+1}(\mathbf{x}',\mathbf{y}) + \mathbb{M}_{e,t}\cdot\rho(\mathbf{x},\mathbf{y})\cdot\max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\} + \\ & \lambda\cdot\mathbb{M}_{e,t}\cdot(1-\rho(\mathbf{x},\mathbf{y}))\cdot\bar{S}_{t+1}(\mathbf{x}',\mathbf{y})\Bigg],\end{aligned} \tag{29}$$

$$\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}) = \frac{\int_{\mathcal{Y}} \mathbb{1}\{\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) < S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\cdot S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\ \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}{\int_{\mathcal{Y}} \mathbb{1}\{\widehat{S}_{t+1}(\mathbf{x}',\mathbf{y}) < S_{t+1}(\mathbf{x}',\tilde{\mathbf{y}})\}\mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}})}.$$

This surplus takes on the same form as (12), but the final continuation value consists solely of a fraction of expected future surplus of the leaving worker. Combining (29), $W_t(\mathbf{x},\mathbf{y}) = \lambda S_t(\mathbf{x},\mathbf{y}) + U_t(\mathbf{x})$, and (25) produces the following wage equation

$$\begin{aligned}w_t(\mathbf{x},\mathbf{y}) = \quad & \lambda f_t(\mathbf{x},\mathbf{y}) + (1-\lambda)c(\mathbf{x},\mathbf{y}) + (1-\lambda)b(\mathbf{x}) + \lambda(1-\lambda)\beta_a(1-\delta)\times \\ & \mathbb{E}_t\Bigg[\mathbb{M}_{u,t}\int_{\mathcal{Y}}\max\{0, S_{t+1}(\mathbf{x}',\mathbf{y})\}\ \mathrm{d}\mathcal{F}_t(\mathbf{y}) - \\ & \mathbb{M}_{e,t}\cdot(1-\rho(\mathbf{x},\mathbf{y}))\bar{S}_{t+1}(\mathbf{x}',\mathbf{y}')\Bigg].\end{aligned} \tag{30}$$

This wage equation mirrors (13), however the final continuation value differs. The sequential auction results in a value that is the convex combination of the competing employers' surpluses.

A.5 Equilibrium Concept

Here, I define the general rational expectations equilibrium and explain the difficulties in solving for it outside of a steady state. I then make the case for a more restrictive but more easily solved partial equilibrium, which I use to take the model to the data. An equilibrium must consist of the solutions to (5), (6), and (9) which characterize equilibrium wages (13) given that free entry assumption drives equilibrium V_t (7) to zero. In general equilibrium, meeting probabilities arise from the measures of employed, unemployed, and vacancies and the matching function. Hence, we add the t subscript to $\mathbb{M}_{u,t}$, $\mathbb{M}_{v,t}$, and $\mathbb{M}_{e,t}$ in all the value functions and include these probabilities in the aggregate state z_t. Now we can define the general rational expectation equilibrium path as follows:

Definition A.1 (Rational Expectations Equilibrium Path)**.** Let u_t be measure of unemployed workers at time t, e_t be measure of employed workers at time t, v_t be the measure of vacancies at time t, ϕ be the on-the-job search effort, and $m : [0,1]^2 \rightarrow [0, \min(u_t + \phi e_t, v_t)]$ be a matching function. Let $\mathcal{W}_t(\mathbf{x}|e) = \int_{\mathcal{Y}} \mathcal{W}_t(\mathbf{x}, \mathbf{y})\mathrm{d}\mathbf{y}$. Given $\{\mathcal{F}_t(\mathbf{y})\}_{t=0}^T$, $\{f_t(\mathbf{x}, \mathbf{y})\}_{t=0}^T$, and initial $\{e_0, u_0, \mathcal{W}_0(\mathbf{x}|u), \mathcal{W}_0(\mathbf{x}|u)\}$, the tuple

$$\{U_t(\mathbf{x}), W_t(\mathbf{x}, \mathbf{y}, \sigma), P_t(\mathbf{x}, \mathbf{y}, \sigma), V_t, w_t(\mathbf{x}, \mathbf{y}, \sigma), \mathcal{W}_t(\mathbf{x}|u), \mathcal{W}_t(\mathbf{x}, \mathbf{y}|e)\}$$

form a rational expectations equilibrium path from time 0 to time T if the following hold.[118]

1. (5), (6), and (9) solve $U_t(\mathbf{x})$, $W_t(\mathbf{x}, \mathbf{y}, \sigma)$, and $P_t(\mathbf{x}, \mathbf{y}, \sigma)$, respectively
2. $w_t(\mathbf{x}, \mathbf{y}, \sigma)$ satisfies (13) for all employed workers
3. $V_t = 0$ at every period t and v_t satisfies (8) [Free Entry]
4. Agents form expectations using $\{\mathcal{F}_t(\mathbf{y})\}_{t=0}^T$ and $\{f_t(\mathbf{x}, \mathbf{y})\}_{t=0}^T$ [Rational Expectations]
5. $\mathbb{M}_{u,t} = \frac{m(v_t, u_t+\phi e_t)}{u_t+\phi e_t}$, $\mathbb{M}_{e,t} = \phi \mathbb{M}_{u,t}$, $\mathbb{M}_{v,t} = \frac{m(v_t, u_t+\phi e_t)}{v_t}$
6. e_t and u_t evolve according to (31) and (33), respectively
7. $\mathcal{W}_t$ evolves by (1) and according to the transitions in (31) and (33)

The main difficulty with for this equilibrium arises from the last condition. The difficulty lies in the fact that this worker distribution ($\mathcal{W}_t$) is endogenous and a part of the state space due to the meeting probabilities ($\mathbb{M}_{e,t}, \mathbb{M}_{u,t}, \mathbb{M}_{v,t}$). $\mathcal{W}_t$ evolves in a complicated way even without human capital evolution. We must track $\mathcal{W}_t$ in order to pin down e_t and u_t and thus v_t from (8) and consequently everything else dependent on the meeting probabilities, $\{U_t, W_t, P_t, V_t, w_t\}$. All these objects must be solved simultaneously, making this equilibrium intractable to solve for over a multidimensional state space. Here I only note the difficulty in finding such an equilibrium if one exists. Establishing a proof of existence or uniqueness of this equilibrium stands as even more challenging. As noted by Menzio and Shi (2011), random search models like the one here remain difficult to solve outside a steady state, because number of employed workers (e_t) and unemployed workers (u_t) depend on the entire distribution of workers across employment states and types.[119] This distribution is not fixed outside of a steady state.

[118] For completeness, $\mathbb{C}_{u,t} = \frac{u_t}{u_t+\phi e_t}$, $\mathbb{C}_{e,t} = \frac{\phi e_t}{u_t+\phi e_t}$.

[119] i_t in Equation 34 exists for accounting purposes, since $e_t + u_t + i_t = N$ where N is the number of agents in the model. There is no population growth, so a new agent fills the place of a dead agent − often referred to as cloning in the search and matching literature.

$$
\begin{aligned}
e_{t+1} = {} & \underbrace{\int_{\mathcal{Y}} \int_{\mathcal{X}|u} (1-\xi_a)\mathbb{M}_{u,t}\mathbb{1}\{S_{t+1}(\mathbf{x},\mathbf{y}) > 0\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|u) +}_{U2E} \\
& \int_{\mathcal{X}|e} (1-\xi_a)(1-\delta)\mathbb{M}_{e,t}e_t\mathrm{d}\mathcal{W}_t(\mathbf{x}|e) + \\
& \int_{\mathcal{X}|e} (1-\xi_a)(1-\delta)(1-\mathbb{M}_{e,t})(1-\omega)e_t\mathrm{d}\mathcal{W}_t(\mathbf{x}|e) + \\
& \int_{\mathcal{X}|e} (1-\xi_a)(1-\delta)(1-\mathbb{M}_{e,t})\omega e_t\mathbb{1}\{S_{t+1}(\mathbf{x},\mathbf{y}) > 0\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|e) - \\
& \Bigg[\underbrace{\int_{\mathcal{X}|e} (1-\xi_a)\delta e_t \ \mathrm{d}\mathcal{W}_t(\mathbf{x}|e) +}_{\text{Exogenous } E2U} \\
& \underbrace{\int_{\mathcal{Y}} \int_{\mathcal{X}|e} (1-\xi_a)(1-\delta)(1-\mathbb{M}_{e,t})\omega e_t\mathbb{1}\{S_{t+1}(\mathbf{x},\mathbf{y}) \leq 0\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|e) +}_{\text{Endogenous } E2U} \\
& \underbrace{\int_{\mathcal{X}|e} \xi_a e_t\mathrm{d}\mathcal{W}_t(\mathbf{x}|e)}_{E2I}\Bigg]
\end{aligned}
\tag{31, 32}
$$

$$
\begin{aligned}
u_{t+1} = {} & \underbrace{\mu_a i_t}_{\text{Entrants}} + \underbrace{\int_{\mathcal{X}|e} (1-\xi_a)\delta e_t \ \mathrm{d}\mathcal{W}_t(\mathbf{x}|e) +}_{\text{Exogenous } E2U} \\
& \underbrace{\int_{\mathcal{Y}} \int_{\mathcal{X}|e} (1-\xi_a)(1-\delta)(1-\mathbb{M}_{e,t})\omega e_t\mathbb{1}\{S_{t+1}(\mathbf{x},\mathbf{y}) \leq 0\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|e) +}_{\text{Endogenous } E2U} \\
& \underbrace{\int_{\mathcal{X}|u} (1-\xi_a)(1-\mathbb{M}_{u,t})u_t\mathrm{d}\mathcal{W}_t(\mathbf{x}|u) + \int_{\mathcal{Y}} \int_{\mathcal{X}|u} (1-\xi_a)\mathbb{M}_{u,t}\mathbb{1}\{S_{t+1}(\mathbf{x},\mathbf{y}) \leq 0\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|u) -}_{U2U} \\
& \Bigg[\underbrace{\int_{\mathcal{Y}} \int_{\mathcal{X}|u} (1-\xi_a)\mathbb{M}_{u,t}\mathbb{1}\{S_{t+1}(\mathbf{x},\mathbf{y}) > 0\}\mathrm{d}\mathcal{F}_t(\mathbf{y})\mathrm{d}\mathcal{W}_t(\mathbf{x}|u) +}_{U2E} \\
& \underbrace{\int_{\mathcal{X}|u} \xi_a u_t\mathrm{d}\mathcal{W}_t(\mathbf{x}|u)}_{U2I}\Bigg]
\end{aligned}
\tag{33}
$$

$$i_{t+1} = \quad (1-\mu_a)i_t + \underbrace{\int\limits_{\mathcal{X}|e} \xi_a e_t \mathrm{d}\mathcal{W}_t(\mathbf{x}|e)}_{E2I} + \underbrace{\int\limits_{\mathcal{X}|u} \xi_a u_t \mathrm{d}\mathcal{W}_t(\mathbf{x}|u)}_{U2I} \tag{34}$$

The directed search literature circumvents this problem, because directed search makes the meeting probabilities independent of the distribution of worker types across employment states and types (Menzio and Shi, 2011; Menzio, Telyukova, and Visschers, 2016). Employers post wages to induce a self-selection of job applicants. Job applicants self-sort and apply in different submarkets, making the meeting probabilities depend only on the number of applicants as all applicants are the same type.

However, this achievement comes at some costs. First, employers attract specific worker types rules out any notion of skill mismatch, which some evidence suggests plays a significant role in wage dispersion (Slonimczyk, 2013). Second, the most recent innovations in directed search models with worker heterogeneity like Menzio, Telyukova, and Visschers (2016) only introduce discrete skills (i.e. age and experience) to my knowledge, which appear ill-equip to handle continuous multidimensional skill like cognitive and manual skills. While seemingly not discussed in the literature, this discreteness appears to contribute significantly to the existence of the block recursive equilibrium. An infinite number of submarkets would need to exist given a continuum of worker types in cognitive and manual skills in order to separate out each multidimensional skill type across submarkets. However, an infinite number of submarkets and no mass points for any one worker type suggests that in the limit there will be only one worker in each submarket queue who is hired with certainty. In this limiting case, it seems directed search implements the outcome from an assignment model with job destruction and productivity shocks, because search frictions do not emerge in submarkets with a continuum of types. Overcoming this drawback likely requires discretizing the skill types, which reenforces the first drawback. Wage differentials due to skill mismatch will be attributed to noise in such a model after collapsing the support of worker types.

Workers and employers face the same distribution of skill requirements in this model, because employer draw skill requirements after meeting the worker. This assumption along with free entry and exogenous meeting probabilities remove the endogenous distribution of worker types or employer types from the state space along a rational expectations equilibrium path. These assumptions eliminate the problem of tracking the endogenous distribution of worker types, however they make the model a partial equilibrium model. While restrictive, these assumptions keep the model tractable while permitting enrichment of the model with multidimensional skills (a necessity to generate secular, non-monotonic changes in the wage distribution). Postel-Vinay and Moscarini (2009) and Robin (2009) also assume exogenous meeting rates to examine labor market dynamics in response to aggregate productivity shocks.

Hawkins and Acemoglu (2014) state that exogenous meeting rates make such a model unsuitable for general equilibrium analysis. In this model, the partial equilibrium misses out on general equilibrium feedback to the meeting rates.[120] Estimating meeting rates ($\mathbb{M}_{e,t}$, $\mathbb{M}_{u,t}$) which change over time may approximate to the general equilibrium solution. However, it is difficult to say how well such a solution approximates the general equilibrium solution without computing the general equilibrium solution. But the general equilibrium solution will also have to generate the same moments (i.e. transition rates) as the partial equilibrium solution to estimate its structural parameters, thus estimating ($\mathbb{M}_{e,t}$, $\mathbb{M}_{u,t}$) to match a target over time may improve

[120] Hawkins and Acemoglu (2014) do not provide any evidence as to how important this feedback is, let alone whether it is important enough to make the partial equilibrium analysis unsuitable for long-run macro level analysis of wages and job selection. This question along with how directed search may resolve this issue are future avenues for research.

the approximation.

A.6 Identification

Provided a sufficiently rich panel data set, we can jointly identify the parameters of the parametric model in Section 4. The following argument only serves to show an identification strategy of the estimated parameters and provide guidance on what moments to target in lieu of the necessary, rich data to implement such a strategy. I target moments carrying much of the same information as the argument ascribes. This identification argument assumes known values for the externally calibrated parameters discussed in Section 4.3 ($\tilde{\beta}$, θ_0, ξ_a) and a given λ. It builds on the argument in Lise and Postel-Vinay (2016) but exploits workers in the terminal period of work life rather than a closed form solution. The data necessary parallels the NLSY panel described in Appendix B.5 but includes workers in the terminal period of work life and their terminal $\mathbf{x}$ as well as gives the reason for a job separation.[121]

Assume we observe initial and terminal $\mathbf{x}$'s (with age) in the data as well as wages without measurement error, $\mathbf{y}_t$, and employment status. Let t_i be the first period a worker i's work life and T_i be the last possible period of a worker's work life. The maximum wage possible for workers age $T_i - t_i + 1$ at time t is $w_t(\mathbf{x}, \mathbf{y}, \sigma) = f_t(\mathbf{y}, \mathbf{y}; \alpha_t) = x_G\Big[\alpha_{0,t} + y_C(\alpha_{C,t} + \alpha_{CC,t}x_C) + y_M(\alpha_{M,t} + \alpha_{MM,t}x_M)\Big]$. Given x_G, wage differentials of maximum wages across $\mathbf{y}$ for worker's age $T_i - t_i + 1$ at time T_i identify $\alpha_{C,t}$, $\alpha_{CC,t}$, $\alpha_{M,t}$, $\alpha_{MM,t}$ at T_i. The level (average) of these maximum wages for worker's age $T_i - t_i + 1$ pins down $\alpha_{0,t}$ conditional on x_G. This argument gives α_t's conditional on x_G. Implicitly, we align the model to the data assuming maximum wages across $\mathbf{y}$ in the model correspond to maximum wages in the data. This imposition and the level of maximum wages also pin down the alignment parameters (ζ_C, ζ_M) conditional on x_G, because $\hat{x}_i^{\zeta_i} = y_i$ at the maximum wage for $i = C, M$.[122]

Wage differentials of identical workers $(x_C, x_M, age, \mathbf{y})$ pin down the θ_1 parameter for the *i.i.d.* random variable ε. Wage differences for such terminal period workers only arise due to ε. Knowing the distribution of ε, wage differentials of workers $(x_C(0), x_M(0), \mathbf{y})$ age t and $t+1$ hired upon entry pin down (γ_1, γ_2). With $(\theta_1, \gamma_1, \gamma_2)$, the distribution of x_G is identified at time t up to some constant γ_0 given the observed $\mathcal{V}_t(\mathbf{x})$. Thus, the maximum wages for workers age $T_i - t_i + 1$ at time t along with wage differentials for entering workers where the unemployment duration approaches zero (i.e. hired upon entry) separately identify α_t and the parameters of x_G upon to some constant γ_0.

Thus, sufficient wage and $(\mathbf{x}, \mathbf{y})$ observations for workers in their initial period provide information to identify $(\gamma_1, \gamma_2, \theta_1)$. While, sufficient wage and $(\mathbf{x}, \mathbf{y})$ observations for workers in their terminal period provide information to identify α_t, ν's, and κ's. Ultimately, wage differentials across $\mathbf{y}$ help determine α_t and conditional wage moments help determine (γ_1, γ_2). I capture these features with changes in wage percentiles, mean wages, wage variance, changes in occupational wages, and the coefficients of age and age^2 in a regression of initial skills and skill requirements. These moments provide information on wage differentials across $\mathbf{y}$. θ_1 affects the dispersion of wages of similarly skilled workers in high skill requirement occupations, thus the right tail of wages serve to capture this information on θ_1.

Given $(\alpha_t, \lambda, \mathbf{x}, \mathbf{y})$, comparisons of wage differentials and employment matches for workers age $T_i - t_i + 1$ hired from unemployment separately identify $(\nu_C, \nu_M, \kappa_C, \kappa_M)$.[123] For example, consider age $T_i - t_i + 1$

[121] The CPS identifies "leavers" and "losers" as the reason for unemployment, referring to voluntary and involuntary unemployment on the part of the worker (IPUMS-CPS, University of Minnesota, 2016).

[122] This argument requires maximum wages across $\mathbf{y}$ correspond to some workers in the terminal period so that we observe $\widehat{\mathbf{x}}$ in the data.

[123] Obviously, we are unlikely to observe workers hired out of unemployment in the period before retirement in the data,

workers hired from unemployment identical in ($\mathbf{y}$, ε, x_M) but not x_C where $\underline{x}_C < y_C$ for some and $x_C = y_C$ for the others. The wage differential (conditional on $\underline{x}_C$ and α_t) between these two groups at time T_i identifies κ_C.[124] Alternatively, we can identify $(\nu_C, \nu_M, \kappa_C, \kappa_M)$ using the set of all observed matches for workers age $T_i - t_1 + 1$. S_t for a worker in the terminal period can be written as

$$S_t(\mathbf{x}, \mathbf{y}) = \quad x_G \Big[\alpha_{0,t} + \alpha_{C,t} y_C + \alpha_{M,t} y_M + \alpha_{CC,t} y_C x_C + \alpha_{MM,t} y_M x_M - b_0 + \\ -\nu_C (x_C - y_C)^2 - \nu (x_M - y_M)^2 + \\ -(\kappa_C - \nu_C) \min\{x_C - y_C, 0\}^2 - (\kappa_M - \nu_M) \min\{x_M - y_M, 0\}^2 \Big].$$

We observe the set of acceptable offers given $\mathbf{x}$ $\{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) \geq 0\}$ and thereby observe its boundary set $\{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) = 0\}$. We observe $\mathbf{x}$ from workers initial and terminal skills. Consider a case where $x_C > y_C$, $x_M = y_M$, and $\mathbf{y} \in \{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) = 0\}$, then we have

$$0 = \alpha_{0,t} + \alpha_{C,t} y_C + \alpha_{M,t} y_M + \alpha_{CC,t} y_C x_C + \alpha_{MM,t} y_M^2 - b - \nu_C (x_C - y_C)^2$$

which identifies ν_C up to the scalar b_0 given α_t. Similar comparisons yield $(\nu_M, \kappa_C, \kappa_M)$ up to scale, hence comparisons of acceptable jobs to workers with similar $\mathbf{x}$ provides information to identify $(\nu_C, \nu_M, \kappa_C, \kappa_M)$. I incorporate this information through the observed cross correlations of $(x_C(0), x_M(0), y_C, y_M)$.

Conditional on the rest of the model parameters, Γ_d is identified by comparing the set of accepted jobs for entering workers with the same starting $\mathbf{x}(0)$ but different initial unemployment spell lengths. Skills depreciate during unemployment spells, thus the job a worker obtains out of unemployment carries information about how fast skill depreciate. Intuitively, skills could not have depreciated to the point where the worker's $\mathbf{x}$ does not generate positive surplus with the employer $\mathbf{y}$. Conditional on the rest of the model's parameters, differences in the set of jobs for initially identical workers come from $\mathbf{x}$, which consists of known $\mathbf{x}(0)$, known unemployment spell duration, and unknown Γ_d. Observing $\{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) = 0\}$, Γ_d is identified conditional on all the other parameters. In practice, I target the average level and dispersion of wage drops following an unemployment spell to estimate the two parameters of Γ_d. Conditional on the other model parameters, Γ_d governs wage drops following an unemployment spells in the same spirit as comparisons of acceptance sets for identical workers.

Identification of Γ_h comes from again comparing workers with similar starting skills but experience different employment-unemployment spell lengths. Given the other model parameters (Δ) and the observed set $\{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) = 0\}$, we can write the surplus function for an entering worker who experiences as unemployment spell one period followed by employment as

$$\begin{aligned} 0 &= f_t(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y}) - b(\mathbf{x}) + \Omega(\mathbf{x}', \mathbf{y}; \Delta), \\ \mathbf{x} &= \mathbf{x}(0) + \ \Gamma_D \cdot \max\{\mathbf{x}(0), \mathbf{0}\}, \\ \mathbf{x}' &= \mathbf{x} + \ \Gamma_H \cdot \max\{\mathbf{y} - \mathbf{x}, \mathbf{0}\} + \ \Gamma_D \cdot \max\{\mathbf{x} - \mathbf{y}, \mathbf{0}\} \end{aligned}$$

where Ω is the continuation value solving backwards to obtain S_{t+1} given model parameters Δ. Given Γ_d,

making the direct application of this strategy impractical. This argument only serves to argue identification of the estimated parameters exists.

[124] Coming out of unemployment wipes the history of workers. Thus knowing λ, it is possible to identify $(\nu_C, \nu_M, \kappa_C, \kappa_M)$ with all wages of workers coming out of unemployment given $\mathbf{x}$. However, such an argument also requires knowledge of other parameters like Γ_d, Γ_h and those of $\mathcal{F}_t(\mathbf{y})$. Using workers in the terminal period eliminates the need to know parameters that enter the continuation value to identify $(\nu_C, \nu_M, \kappa_C, \kappa_M)$.

the only unknown is Γ_h which is identified up to scale with the observed set $\{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) = 0\}$. Given Γ_h and Γ_d, the sequence $\{\mathbf{x}(t)\}_{t=t_i}^{T_i}$ can be identified for each worker based on (17).[125]

Conditional on the other parameters, $\mathcal{F}_t(\mathbf{y})$ is identified over the union of all sets where $\mathbf{y}$ is acceptable to an $\mathbf{x}$, i.e. $\bigcup_{\mathcal{Y}} \{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) \geq 0\}$. All potential employers draw skill requirements from the same distribution independently, however changes in skill requirements do not map one-to-one to changes in employment shares over $\bigcup_{\mathcal{Y}} \{\mathbf{y} : S_t(\mathbf{x}, \mathbf{y}) \geq 0\}$. Nonetheless, employment shares across $\mathbf{y}$ map out $\mathcal{F}_t(\mathbf{y})$ given the other model parameters that define S_t. Thus, I target changes in employment shares for occupational group in practice.

All endogenous separations are mutual through the lens of the model and result from changes in $\mathcal{F}_t(\mathbf{y})$ and a permanent productivity shock (ω). Thus, δ shocks create involuntary unemployment whereas ω shocks may result in voluntary unemployment. Thus, the average ratio of voluntary to involuntary unemployment and the average employment-to-unemployment (E2U) transition rate identify δ and ω given $\mathcal{F}_t(\mathbf{y})$. The unemployment-to-employment (U2E) transition rate at time t identifies $\mathbb{M}_{u,t}$ given all other parameters. The employment-to-employment (E2E) transition rate at time t identifies $\mathbb{M}_{e,t}$ given all other parameters.

Finally, we can solve backwards and write the wage continuation value as a function Ω of b_0 given all other parameters Δ. Wages out of unemployment then identify b_0 given Δ and $\mathbf{x}$ as shown in (35).

$$(1-\lambda)b(\mathbf{x}) + \Omega(b_0; \Delta) = w_t(\mathbf{x}, \mathbf{y}, \lambda; \Delta) - \lambda f_t(\mathbf{x}, \mathbf{y}; \Delta) + (1-\lambda)c(\mathbf{x}, \mathbf{y}; \Delta). \tag{35}$$

Thus, b_0 is identified up to scale conditional on all other parameters. This completes the argument for joint identification of the parameters. As mentioned, I target moments related to the information contained in such an identification strategy even though the data does not permit its full implementation.

B Data Appendix

B.1 Current Population Survey (1979-2010)

I use the Current Population Survey's Outgoing Rotation Group (CPS ORG), because of its timespan, informational content, frequency, and comparability over time. These features make it more appropriate for my use than other nationally representative surveys like the SIPP, PSID, or SCF. The CPS ORG provides monthly data from as far back as 1979 and covers every year up to 2016 (National Bureau of Economic Research, 2016). I make use of the CEPR Uniform Data Extracts for the CPS ORG (Center for Economic and Policy Research, 2017). The CEPR constructs monthly extracts from the NBER Merged ORG extracts from 1979 to 1993 and the CPS Basic data from 1994 onwards. I use these extracts from CEPR and their publicly available programs to construct a consistent, monthly dataset from 1979 to 2010 of the CPS Outgoing Rotation Group year by year. These extracts contain monthly cross-sectional data on earnings, employment status, occupation and industry codes, age, educational attainment, gender, and self-employment status among other variables. These CPS ORG extracts contain about 25,000 records each month before merging with the occupational skill scores and imposing sample restrictions, which I describe later.

[125]

$$\begin{aligned}\Omega(\mathbf{x}', \mathbf{y}; \Delta) = \beta_a(1-\delta)\Big[&-\lambda \mathbb{M}_{u,t} \int_{\mathcal{Y}} \max\{0, S_{t+1}(\mathbf{x}', \tilde{\mathbf{y}})\} \, \mathrm{d}\mathcal{F}_t(\tilde{\mathbf{y}}) + \\ &(1-\mathbb{M}_{e,t})\tilde{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \mathbb{M}_{e,t} \cdot \rho(\mathbf{x}, \mathbf{y}) \cdot \max\{0, S_{t+1}(\mathbf{x}', \mathbf{y})\} + \\ &\mathbb{M}_{e,t} \cdot (1-\rho(\mathbf{x}, \mathbf{y})) \cdot \Big[\widehat{S}_{t+1}(\mathbf{x}', \mathbf{y}) + \lambda\big(\bar{S}_{t+1}(\mathbf{x}', \mathbf{y}) - \widehat{S}_{t+1}(\mathbf{x}', \mathbf{y})\big)\Big]\Big].\end{aligned}$$

B.1.1 Wage Measurement, Top-Coding, and Imputation

Schmitt (2003) provides a detailed discussion of issues related to measuring hourly wages with the CPS ORG. I summarize the main issues here with respect to wage measurement, top-coding (commonly known as censoring), and imputation.

The CPS ORG wage records arguably provide a more accurate wage measure than the CPS March Supplement as they measure most wages (approximately 60%) at a point in time (Mishel, Bivens, Gould, and Shierholz, 2012; Lemieux, 2006).[126] For consistency purposes, I exclude overtime, tips, and commission (otc) from hourly wage records. The complicated nature of when and how the CPS reports this compensation makes it intractable to create a sensible series including otc for these records over more than a few years as noted in Mishel, Bivens, Gould, and Shierholz (2012) and Schmitt (2003).[127] The remaining 40% of ORG records report a constructed measure of hourly wages using weekly earnings and usual hours worked per week, which includes otc by the construction of weekly earnings. This measurement contains substantially more measurement error compared to the point in time measure (Lemieux, 2006). The March CPS permits only a constructed measure of hourly or weekly earnings from total earnings, weeks worked, and usual hours worked each week. Consequently, the measurement error in the March CPS wages seems significantly higher than the ORG as documented by Lemieux (2006). Hence, ORG wage records arguably provide a more accurate wage measure even though they do not measure all wages at a point in time.

Each year, 1-3% of these 40% of ORG records exceed the top-code threshold except in the 1980s where the share grows due to nominal earnings growth with no increase in the top-code threshold. Following Schmitt (2003) with the CEPR programs, I impute these top-coded weekly earnings using a log-normal imputation. The imputation estimates the mean of the wage distribution by gender above the top-code threshold and replaces the top-coded wage with this value. The log-normal imputation procedure provides for a smoother wage series over time in terms of mean and variance than the commonly used Pareto imputation (Schmitt, 2003). As seen in Schmitt (2003), these top-coded records have little impact on wage percentiles − a key measurement of interest here − compared to the wage mean and variance.

B.1.2 Occupational Code Harmonization

The CPS employs the Census occupational coding structure which is derived from the Standard Occupational Classification (SOC) and the North American Industry Classification System (NAICS). Major occupational coding changes in 1983 and 2003 and a minor change in 1992 complicate the construction of a consistent set of occupational code from 1979 to 2010. These changes introduces discontinuities in employment shares and average wages by occupation over time. The coding change in 1983 introduces 64 new occupational titles while the 2003 coding change reduces the number of titles by collapsing and expanding occupational categories.[128] Dorn (2009) provides a crosswalk to create a balanced panel of occupations from 1983 onward. This balanced panel consists of aggregated occupational categories shown in his Appendix Table 2. I use this *occ1990dd* crosswalk to harmonize the occupational titles from 1979 to 2010, which results in 246 occupations on which to construct DOT/ONET scores. From 1979 to 1983, occupational employment shares and average wages cannot be constructed for 64 occupations, because they only begin to appear in 1983. These occupational titles range from human resource managers to occupational therapists to locksmiths to

[126] These hourly wage records also rarely cross the top-coding threshold of 99.99, so I follow Schmitt (2003) and make no top-coding adjustment on them.

[127] In many cases, reconstructing hourly wages from weekly earnings in order to include otc for these records produces hourly wages that imply otc is counterfactually negative.

[128] See Dorn (2009) Appendix Table 1.

machine feeders. Discontinuities persist in occupational series (e.g. cognitive skills, employment shares) with this harmonized set of occupational titles. I apply the method of Mishel, Schmitt, and Shierholz (2013) and smooth any occupation related series at the major coding break years 1983 and 2003. This adjustment produces series similar to the original series overall but with slight differences. For example, Figure 26 shows the slight magnitude differences in employment share and occupational wage changes. The only patterns change comes from average low-skilled wages falling in the 2000s while still maintaining their relative distance from medium and high-skilled average wages.

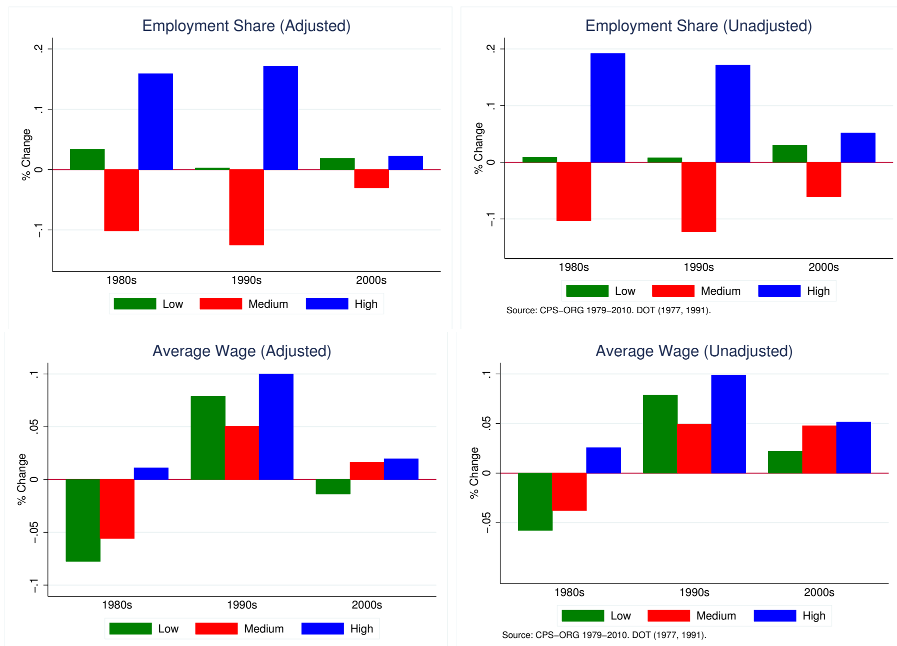


Figure 26: Occupational Coding Break Adjustment

B.2 Dictionary of Occupational Titles (1977, 1991)

The Dictionary of Occupational Titles (DOT) provides measures related to the job requirements for 12,099 occupational titles (U.S. Department of Labor, 1991). Job analysis reports serve as the source of the measures, and these reports come from combinations of on-site observation, interviews, and external information (e.g. information from trade associations) (Yamaguchi, 2012). Job analysis measures the "worker attributes that contribute to successful job performance" (U.S. Department of Labor, 1991). However, the DOT measures these attributes based on tasks the worker performs rather than the worker skills. This distinction along with the use of external information arguably justifies treating DOT measures as constructed independently of the workers' skills at the time of measurement. Thus, these measures allow me to construct cognitive and manual measure analogous to $\mathbf{y}$ in the model. In contrast, DOT's modern replacement O*NET (Occupational Information Network) collects data directly from incumbent workers, making it more difficult to argue independence of worker skills at the point of measurement (O*NET, 2016). The DOT is also updated over the time period of interest as a Revised Fourth Edition emerged in 1991 in addition to the Fourth Edition in 1977 whereas O*NET provides no time variation in task measurements over the period considered.[129]

[129]O*NET provides waves for only 2008 and 2013 as of now. I explore the possibility of mapping the DOT to O*NET over time using job attributes that appear in both datasets. I conclude that the differences between the DOT and O*NET are too vast to permit a full, consistent mapping of task measures across these two datasets.

Of course, the DOT is not without its own shortcomings − many of which O*NET aims to improve upon. Miller, Treiman, Cain, and Roos (1980) provide a critical review of the DOT. These criticisms include the limited time dimension of DOT updates and its outmoded nature with respect to new occupations. The occupation coding change in 2003 presents a challenge for using the DOT beyond 2000 as it introduces new occupational titles. Yamaguchi (2012) drops observations beyond the year 2000 for this reason. However, this work aims to understand wage and occupational structure changes up to 2010. Hence, I construct DOT measures for these new occupations using weighted combinations of older but similar occupational titles and validate these imputed measures with O*NET measures. These new occupational titles consists mainly of informational and technology occupations like computer support specialists and computer software engineers.

The ICPSR distributes DOT measures for the 1980 and 1990 Census occupational codes for DOT (1977) and DOT (1991), respectively. England and Kilbourne (1980) and the U.S. Department of Labor, U.S. Employment Service, and the North Carolina Occupational Analysis Field Center (1991) produce these DOT measures by aggregating the 12,099 occupational titles of the DOT to the Census occupation level. To do so, they make use of the April 1971 CPS Monthly File and the so-called Treiman file, which ultimately record a sample of the Census in DOT and Census 1970 and 1980 occupational titles.[130] Using this matching file, they take weighted averages of the DOT measures to aggregate to the Census occupational level. I match these Census occupational codes to the *occ1990dd* harmonized occupational code and aggregate again, taking weighted averages of the DOT measure to reach the *occ1990dd* level. I use the respective Census weights for this procedure. This procedure compresses the variance in these measures, thus it likely leads to underestimating the true level of dispersion among the DOT task measures. However, the literature commonly uses such averaging to aggregate the DOT or O*NET to a level to merge with the NLSY or CPS (Dorn, 2009; Acemoglu and Autor, 2011; Yamaguchi, 2012; Lise and Postel-Vinay, 2016).[131] I retain the measures from the DOT shown in Table 13 and a measure of the physical strength a job requires.

Table 13: DOT Task Complexity Measures

	Name	Ability
G	General Learning Ability	Learn, reason, and make judgments
V	Verbal Ability	Understand use words effectively
N	Numerical Ability	Understand and perform mathematical functions
S	Spatial Ability	Visualize three dimensional objects from two dimensions
P	Form Perception	Perceive and distinguish graph detail
Q	Clerical Perception	See and distinguish verbal details
K	Motor Coordination	Coordinate eyes, hands, fingers
F	Finger Dexterity	Finger and manipulate small objects
M	Manual Dexterity	Handle placing and turning motions
E	Eye-Hand-Foot Coordination	Motor responsiveness to visual stimuli
C	Color Discrimination	Match and discriminate colors

Source: U.S. Department of Labor (1991).

These measures range from 1 to 5 where 1 indicates the most complex usage of the ability and 5 indicates the least complex.

[130] Autor, Levy, and Murnane (2003) describe this procedure thoroughly in their Section A.2.

[131] Sanders (2016) puts all weight on the disaggregated occupation with the highest employment share when aggregating up to the Census level to merge the NLSY and O*NET.

B.3 CPS-DOT Construction and Sample Restrictions

I construct an annual CPS-DOT dataset to analyze wage and employment share trends. I impose some restrictions on the data, which can be followed in Table 14. First, I restrict the sample to the population aged 18 to 65. Second, I restrict the sample to include only observations with a valid wage and occupational code. This restriction eliminates all unemployed workers and workers out of the labor force. Third, I merge the DOT to the CPS based on the harmonized *occ1990dd* occupational code. I impute occupations with missing scores using weighted average DOT scores from similar occupations based on occupational descriptions. For example, I impute the DOT measures for occupation "secretary (not specific)" using all other types of secretaries.[132] I drop some observations after merging in the DOT or O*NET due to dropping armed forces members and unpaid family farm workers. Finally, I keep all non-self-employed workers aged 18 to 64, and I follow the literature in eliminating implausibly low or high values by dropping wage records below \$1 or above \$100 in 1989 terms (Lemieux, 2006). I show the remaining number of valid cases per annum in the last column of Table 14. Table 16 present demographic and distributional statistics for the sample at the start and end years.

B.4 O*NET Comparability

I use the same procedure described for DOT to construct a CPS-O*NET dataset. O*NET lists occupations using the Standard Occupation Classification (SOC) system. I use a crosswalk from National Crosswalk Service Center (2016) to map SOC codes from 2000 and 2010 into Census occupation codes and hence the *occ1990dd* harmonized code.[133] I select measures from O*NET that align with the DOT measurements based on their descriptions. In some cases, multiple O*NET measures correspond to the DOT measure. For instance, O*NET element IDs 1A1a1-1A1a4 correspond to verbal ability. In other cases, a single O*NET measure corresponds to the DOT measure like manual (M) and finger (F) dexterity and color discrimination (C). As described in 3.2, I use the first principle component of general learning ability, verbal ability, and numerical ability measures to construct the cognitive skill score weighted using the CPS ORG weight. I use the first principle component of the other measures (S, P, Q K, M, F, E, C, Strength) to construct the manual skill score. I then linearly rescale the scores into the interval $[0, 1]$. The U.S. Department of Labor updated the DOT in 1991, and O*NET replaced it in the 2000s. I show the DOT and O*NET cognitive and manual skill scores at all occupations with both scores during the 1990s decade of transition in Figure 27.

[132] 90% of all missing scores come from this one occupational title.

[133] I modify the crosswalk manually like Sanders (2016) and impute some O*NET measures as some SOC codes correspond to multiple Census codes and vice versa. Approximately, 70% of the codes map one-to-one.

Table 14: Sample Size Post-Restrictions

Year	Ages 18-65	Valid Wage/Occupation	DOT/O*NET	Additional Restrictions
1979	266,575	161,561	161,561	160,648
1980	313,645	188,230	188,230	187,097
1981	295,931	176,963	176,963	176,031
1982	285,736	167,249	167,249	166,322
1983	283,371	165,764	165,653	164,598
1984	279,684	168,976	168,878	167,839
1985	279,892	172,193	172,086	171,046
1986	273,846	170,856	170,757	169,673
1987	272,186	171,887	171,780	170,693
1988	258,132	164,745	164,647	163,629
1989	262,498	168,233	168,122	167,308
1990	276,736	176,903	176,769	175,820
1991	273,160	171,936	171,797	170,900
1992	268,355	169,499	169,484	168,702
1993	264,119	167,325	167,304	166,438
1994	256,178	162,647	162,623	161,749
1995	252,855	162,280	162,265	161,409
1996	223,258	144,821	144,820	144,070
1997	225,572	147,579	147,579	146,857
1998	225,754	149,332	149,332	148,563
1999	227,599	151,478	151,478	150,783
2000	229,056	153,224	153,224	152,441
2001	244,931	163,121	163,121	162,174
2002	266,531	175,260	175,260	174,243
2003	265,775	172,124	172,124	171,090
2004	261,571	169,246	169,246	168,189
2005	261,116	170,297	170,297	169,159
2006	258,747	169,606	169,606	168,540
2007	256,367	167,882	167,882	166,600
2008	255,574	165,984	165,984	164,588
2009	258,110	161,110	161,110	159,837
2010	257,936	159,431	159,431	158,209

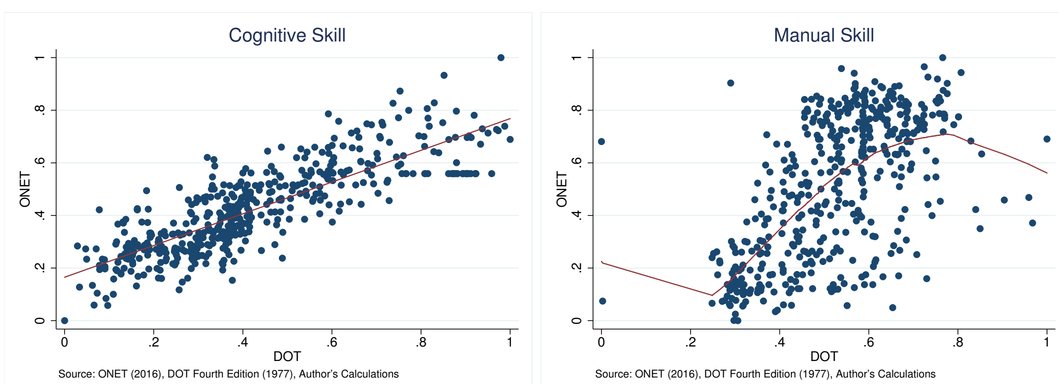


Figure 27: DOT to O*NET from 1992 to 2002

The DOT and O*NET cognitive scores line up well. It is not unreasonable to allow the O*NET score to be an affine transformation of the DOT score as shown by the red line in the left panel of Figure 27.

Table 15: CPS-DOT Summary Statistics

	1979	2010
Age Shares (%)		
Age 18-24	22.29	12.95
Age 25-34	28.84	23.98
Age 35-44	20.03	23.18
Age 45-54	17.11	24.47
Age 55-64	11.73	15.42
Education Attainment (%)		
Less than High School	19.94	7.48
High School Diploma	38.62	29.37
Some college	22.82	30.08
College	12.94	21.98
Advanced	5.67	11.08
Female Share (%)	43.44	48.71
Occupation (%)		
Management & Professional	27.83	40.27
Administrative & Retail Sales	25.06	22.14
Low-Skill Services	12.02	16.32
Production & Craft	4.46	2.64
Operators, Assemblers & Inspectors	11.99	3.81
Transportation, Construction, & Mining	18.65	14.82
Distribution of $\mathbf{y}$		
Mean of y_C	0.388	0.436
Mean of y_M	0.445	0.413
Standard Deviation of y_C	0.206	0.208
Standard Deviation of y_M	0.143	0.150
Correlation (y_C, y_M)	-0.017	-0.111
Mean of Log Wage	2.810	2.906
Variance of Log Wage	0.261	0.376
Sample Size	160,648	158,209

However, the manual scores do not line up well. The noise introduced by numerous O*NET measures corresponding to the (S, P, Q K, M, F, E, C, Strength) measures accounts for some of this difference. In additional, improvements made to measurement on these task aptitudes and possible changes in task content within occupational titles over time account for some of this difference. However, the data does not permit us to distinguish these three items even with identical occupational titles and their DOT and O*NET measures. From this exercise, I conclude that mapping DOT to O*NET appears only reasonable in the case of a limited set of task measurements – in particular the cognitive measurement and measurements that correspond exactly (e.g. manual and finger dexterity). Thus, they do not permit a full mapping across time. A trade off exists between losing information and losing consistency over time. For the reasons described in B.2, I use the DOT for estimation of the model.

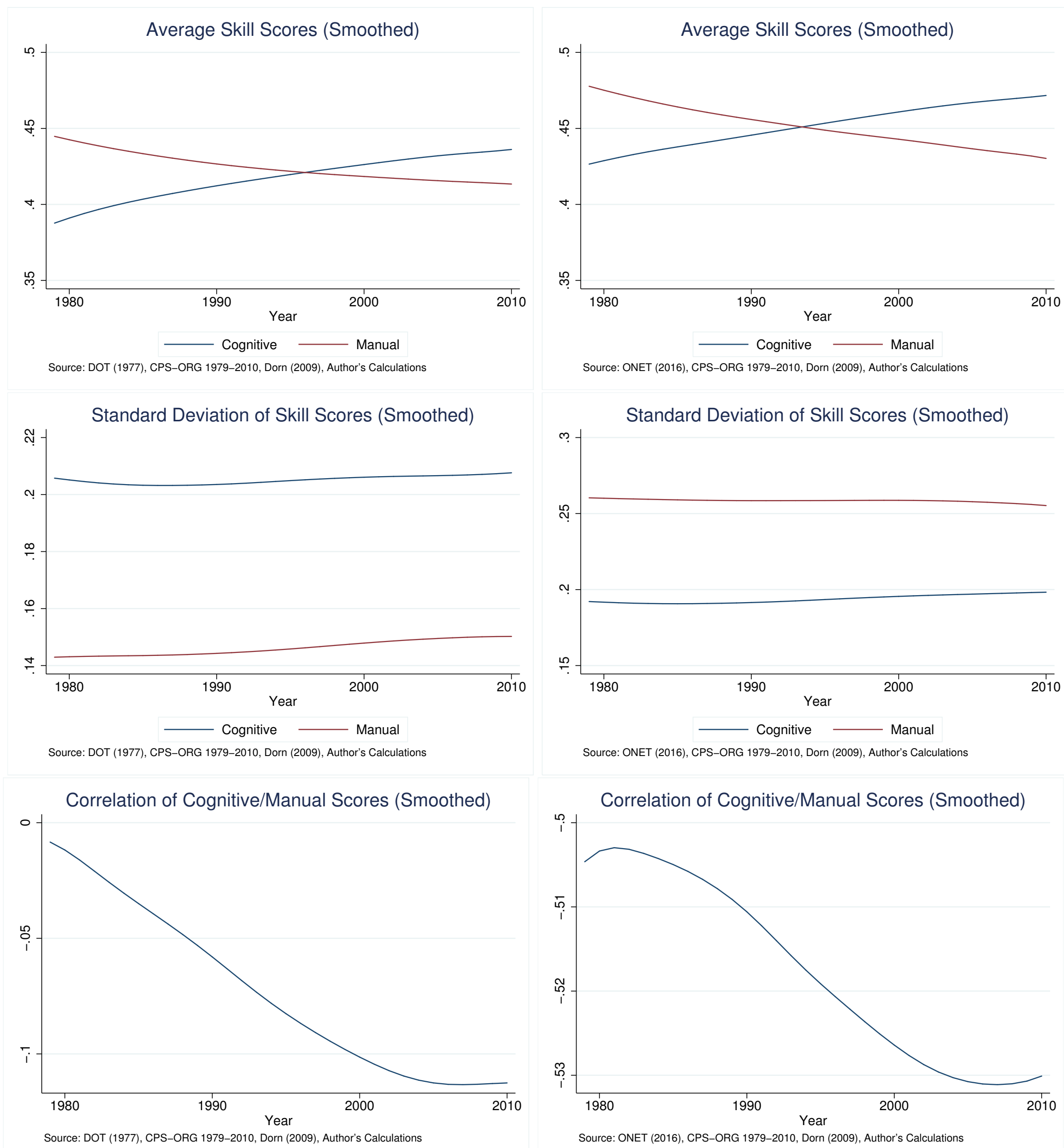


Figure 28: DOT v. O*NET (1979 to 2010)

Constructing cognitive and manual skill requirement scores from DOT versus O*NET results in a different distribution of equilibrium $\mathbf{y}$. However, the main differences occur in the levels and not the evolution of the series as Figure 28 shows.[134] The left panel of Figure 28 indicates moments of $\mathbf{y}$ according to the DOT task measures and the right panel indicates those same moments according to analogous O*NET measures. The mean of cognitive (manual) skills increases (decreases), and cognitive skills become more dispersed over time for both. They also agree as to the decelerating decline in the correlation between cognitive and manual skills. However, they contradict in terms of whether manual skills become more dispersed over time. This difference comes at no surprise given the right panel of Figure 27.

B.5 National Longitudinal Survey of Youth (1979, 1997)

The National Longitudinal Survey of Youth (NLSY) provides ability measures analogous to the DOT task complexity measures. These measures provide a means to construct $\mathcal{V}_t(\mathbf{x})$ and examine the joint distribution of $(\mathbf{x}(0), \mathbf{y})$. The panel and national representative features of the NLSY also provide a means to estimate other data features like the average fall in wages following an unemployment spell. The 1979 cohort consists of 12,686 males and females (Bureau of Labor Statistics, U.S. Department of Labor, 2014a), while the 1997 cohort consists of 8,984 (Bureau of Labor Statistics, U.S. Department of Labor, 2014b). Around half of the observations in each cohort come from an oversample of blacks, Hispanics, and non-black/non-Hispanic economically disadvantaged youth. I drop these respondents, leaving 8,998 and 7,127 respondents for the NLSY79 and NLSY97, respectively.

B.5.1 Construction of $\mathcal{V}_t(\mathbf{x})$

As described in 3.3, I use the Armed Services Vocational Aptitude Battery (ASVAB) test scores to construct $\mathbf{x}(0)$. The ASVAB test consists of scores for mathematics knowledge, arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, general science, coding speed, auto and shop information, mechanical comprehension, and electronics information (Bureau of Labor Statistics, U.S. Department of Labor, 2014a,b). Raw scores between NLSY79 and NLSY97 are not readily comparable for two reasons. First, NLSY79 respondents did a pencil and paper test whereas their 97 counterparts did a computerized test. Segall (1997) accounts for this difference and provides comparative ASVAB scores and weights which I use. Second, the two cohorts took the exam at different ages. NLSY79 took the exam from aged 15 to 22 while NLSY97 took the exam aged 12 to 17. I follow Altonji, Bharadwaj, and Lange (2012) and do a percentile based age mapping of the Segall scores to make the two cohort scores comparable.

Taking these transformed scores, I extract the first two principle components of all the ASVAB scores, and rotate them using the two restrictions on the loading matrix. I restrict mathematical knowledge to load only on cognitive skills, and I restrict auto and shop information to load only on manual skills. Then I linearly rescale these rotated scores into the interval $[0, 1]$ to form $(\tilde{x}_C(0), \tilde{x}_M(0))$. I employ principle component analysis here instead of separating the measures into categories, because some of the ASVAB measures do not categorize as easily as the DOT or O*NET measures like electronics information. This $\tilde{\mathbf{x}}(0)$ does not necessarily align with the estimate $\mathbf{y}$. I perform the following steps to align $\tilde{\mathbf{x}}(0)$ with $\mathbf{y}$. First, I merge $\mathbf{y}$ from the first recorded occupation for the 1979 respondents. Next, I run a log-log regression aimed at minimizing the discrepancy between initial skills and initial job requirements. This step normalizes the level of potential skill mismatch (i.e. the gap between worker skill and job skill requirements). Then, I take the fitted values of $\tilde{\mathbf{x}}(0)$ − call them $\widehat{\mathbf{x}}(0)$ − and use them to construct $\mathcal{V}(\mathbf{x})$ for the two cohorts.[135] I show

[134] I smooth the time series of the moments to reduce sampling noise using Lowess with the optimal bandwidth.

[135] I also allow a transformation of $\widehat{\mathbf{x}}(0)$ into $\mathbf{x}(0)$ in the estimation to better align it with $\mathbf{y}$.

the marginal distributions for this $\widehat{\mathcal{V}}_{1979}(\mathbf{x})$ and $\widehat{\mathcal{V}}_{1997}(\mathbf{x})$ in Figure 29.[136]

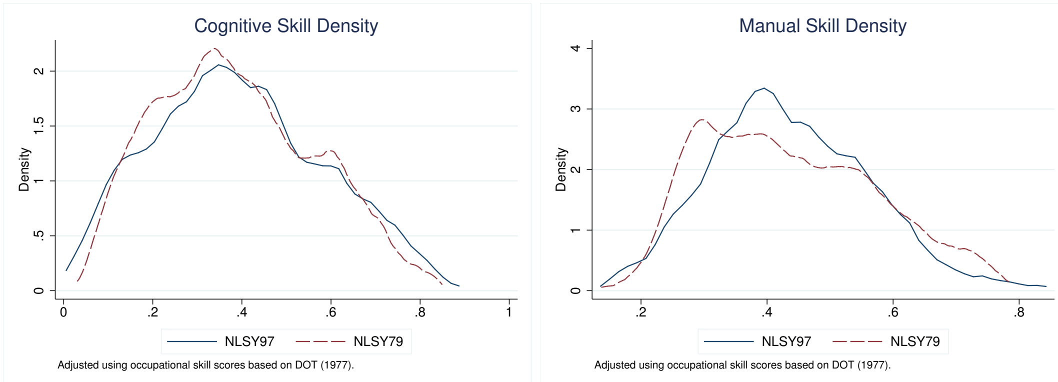


Figure 29: NLSY79 v. NLSY97 Marginal Distributions

The comparable distributions of initial skill show some small changes between the two cohorts. However, the striking similarity of $\mathbf{x}(0)$ across cohorts also suggests that it remains reasonable to treat the distribution of $\mathbf{x}(0)$ as fixed given educational attainment and gender shares. This result is not surprising given that Boehm (2017) uses a similar approach to measure his $\mathbf{x}(0)$ and finds little change in the correlation structure of skill scores between the two cohorts. Thus, $\widehat{\mathcal{V}}_{1979}(\mathbf{x})$ forms the basis for $\mathcal{V}_0(\mathbf{x})$. I reweigh $\widehat{\mathcal{V}}_{1979}(\mathbf{x})$ to reflect changing educational attainment and rising female labor force participation to obtain $\mathcal{V}_t(\mathbf{x})$ over time. This approach remains sensible only if the distribution of cognitive and manual skill remains similar within education-gender cells of cohorts. Figure 30 shows that this appear to be the case in terms of gender. The marginal distributions for females look similar between the two cohorts although manual skills appear to skew more positively for the 1997 cohort. The two cohorts also appear similar with respect to the marginal distribution given an education level. For example, the comparable marginal distributions for cognitive skills at different education attainment levels look similar after adjusting for the time of the ASVAB test shown in Figure 31.

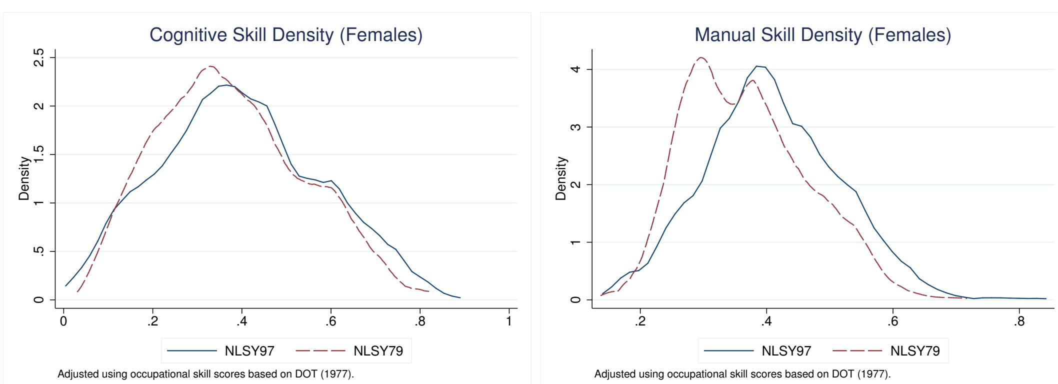


Figure 30: NLSY79 v. NLSY97 Marginal Distributions for Females

[136] I use an Epanechnikov kernel with the optimal bandwidth selected for Figure 29.

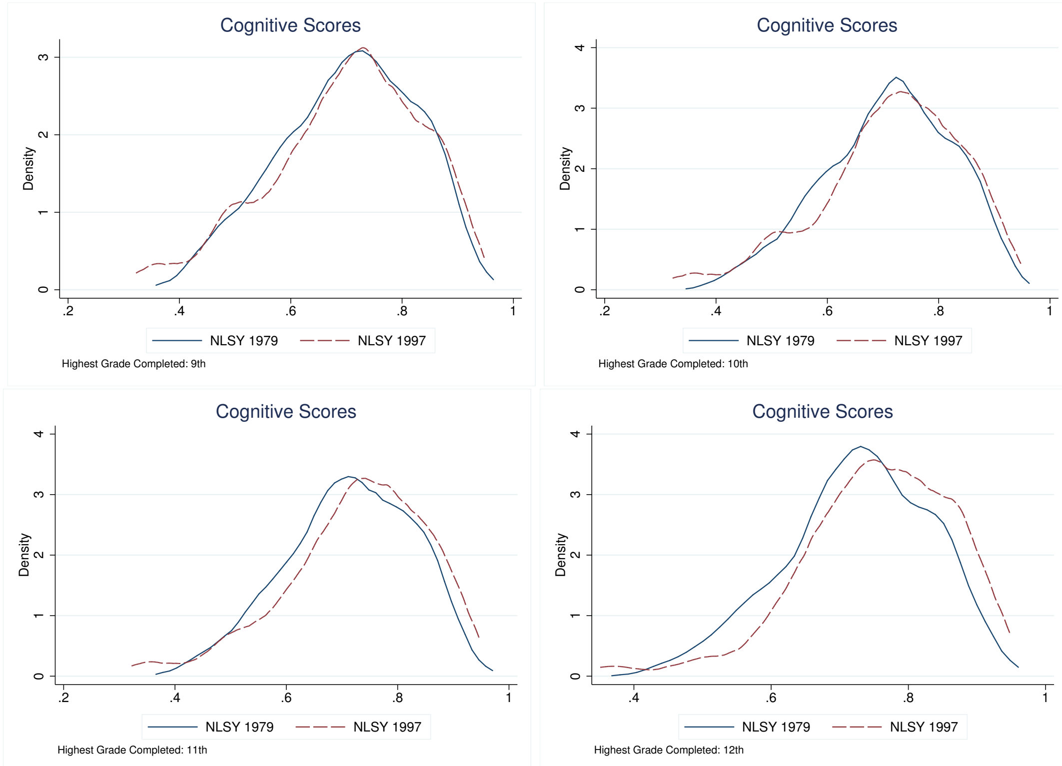


Figure 31: NLSY79 v. NLSY97 Distribution of Cognitive Skills by Highest Grade Completed

My estimate of $\mathcal{V}_t(\mathbf{x})$ only reflects changes in the initial skill distribution due to changes between shares of education and gender groups rather than changes in the distribution of skill within gender and education groups. Thus, $\mathcal{V}_t(\mathbf{x})$ amplifies the initial manual skill bias between males and females shown in Figure 32 as the share of females rises as shown in Figure 34. It also yields an increase in worker cognitive skills shown in Figure 33 as education attainment rises as shown in Figure 34. Comparing NLSY cohorts shows that changes in initial skill within these groups (Figures 30, 31) appear less dramatic than changes in the shares of these groups (Figure 34). This evidence suggests my construction of $\mathcal{V}_t(\mathbf{x})$, holding the within group distribution fixed, reasonable. However, we need more cohorts to definitively argue for this restriction, which are unavailable at this time.

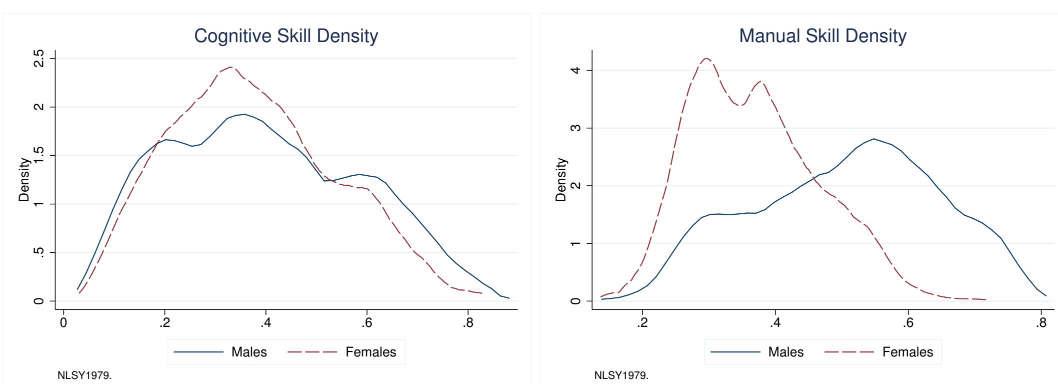


Figure 32: NLSY1979 Initial Skills by Gender

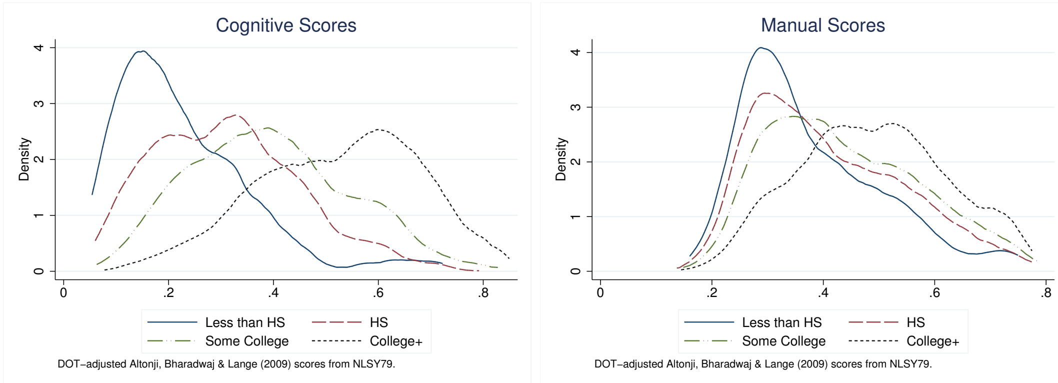


Figure 33: NLSY1979 Initial Skills by Education

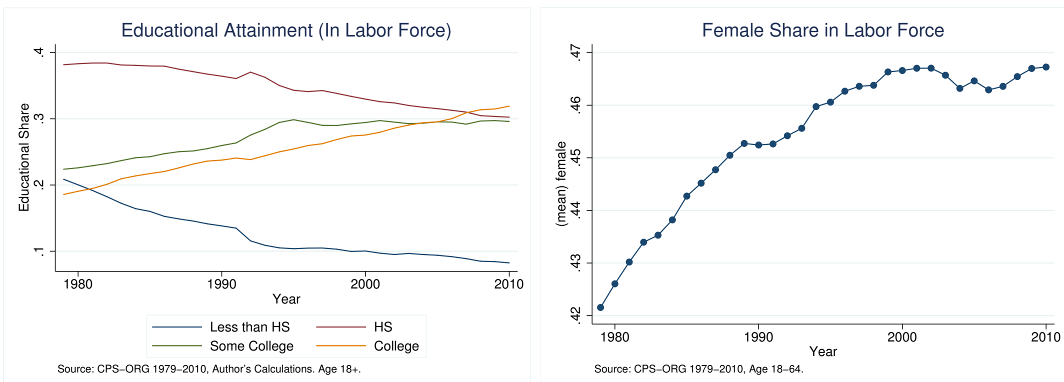


Figure 34: Educational Attainment and Female Share in Labor Force

B.5.2 Construction of Monthly Panel and Sample Restrictions

I construct a monthly panel of workers from the NLSY79 job array. The job array reports the weekly start and end dates of job spells and identifying job numbers. I merge in the corresponding wages, Census occupational codes, and usual hours worked associated with the job numbers. I also merge in demographic data, including gender, race, age, years of schooling, and highest grade completed. I drop oversampled black, Hispanic, and economically disadvantaged workers. I convert monthly-level wages to real 2014 dollar wages using the CPI-U-RS series to make NLSY wages comparable to the CPS-DOT wages. I impute top-coded wages using the same method described in B.3, and trim wages below \$1 or above \$100 in 1989 terms. For workers with multiple jobs within a month, I select the job with the highest earnings that month. I merge in workers' initial cognitive and manual skills along with DOT and O*NET job skill requirements constructed in the previous section. Due to accelerating attrition, I limit the panel to cover only up to 1993 as discussed in Lise and Postel-Vinay (2016). Finally, I reconstruct the sampling weight as in Boehm (2017) and Altonji, Bharadwaj, and Lange (2012) to produce a final weight accounting for attrition, missing ASVAB scores, and hours worked at the job. This process results in a monthly panel of 5,747 male and female workers from 1979 to 1993. Table 16 presents summary statistics of the sample.

Table 16: NLSY79-DOT Summary Statistics

	1979-1993	
Female Share (%)	52.20	
Educational Attainment (%)		
Less than High School	8.49	
High School Diploma	34.37	
Some college	25.06	
College+	32.08	
Occupation (%)		
Management & Professional	29.89	
Administrative & Retail Sales	22.71	
Low-Skill Services	16.43	
Production & Craft	3.15	
Operators, Assemblers & Inspectors	8.16	
Transportation, Construction, & Mining	19.65	
Distribution of $\widehat{\mathbf{x}}(0)$		
Mean of $\widehat{x}_C(0)$	0.439	
Mean of $\widehat{x}_M(0)$	0.611	
Standard Deviation of $\widehat{x}_C(0)$	0.191	
Standard Deviation of $\widehat{x}_M(0)$	0.129	
Correlation of $(\widehat{x}_C(0), \widehat{x}_M(0))$	0.427	
Joint Distribution of $(\mathbf{y}, \widehat{\mathbf{x}}(0))$		
Correlation of $(y_C, \widehat{x}_C(0))$	0.354	
Correlation of $(y_M, \widehat{x}_M(0))$	0.080	
Correlation of $(y_C, \widehat{x}_M(0))$	0.122	
Correlation of $(y_M, \widehat{x}_C(0))$	-0.041	
	1979	1993
Mean of Log Wage	2.462	2.889
Variance of Log Wage	0.148	0.338
Sample Size (Individuals)	5,747	

B.6 Occupational Wage and Employment Changes

The literature commonly presents job polarization as changes in employment shares across the occupational skill distribution. Authors typically rank disaggregated occupations by their mean wage, median wage or education attainment rather than grouping occupations into major categories. They then plot smoothed changes in employment shares across these ranks. A U-shape plot rising at the bottom and the top reveals job polarization either in absolute or relative terms. Absolute means that the bottom and top employment shares rise. Relative means either the top or the bottom employment shares rise, while the middle-skill employment shares shrink the most.

Mishel, Schmitt, and Shierholz (2013) and Lefter and Sand (2011) analyze evidence regarding job and polarization. Both conclude that the evidence fails to some extent to support the narrative of routine-biased technical change as put forward by Autor, Levy, and Murnane (2003) and developed in Acemoglu and Autor (2011) and Boehm (2017) among others. Their critique centers on two pieces of evidence. First, job polarization appears to show similar patterns in the 1980s and 1990s. Thus, factor driving job polarization seem unlikely candidates to account for the abrupt switch from expansion to contraction in the lower half of the wage distribution from the 1980s to the 1990s. Second, the weak correlation between occupational wages and employment shares does not intuitively support a demand-driven story of technological change. Lefter and Sand (2011) use the CPS March supplement and the decadal Census. Mishel, Schmitt, and Shierholz (2013) use the CPS ORG as I do.

I replicate and extend their figures to 2010, ranking occupation using 1979 average wages. Like Mishel, Schmitt, and Shierholz (2013), I smooth over occupational breaks in 1983 and 2003, replacing the wage and employment share changes for each occupation those years with the average change two years before and after the coding break. I exclude farmers and other small sized occupations like wall paper hangers. I extrapolate employment shares and average wages in 1979 for the new 64 occupations that appear in 1983 in order to rank them.[137] I extrapolate using a fractional polynomial time trend. I validate this procedure by extrapolating employment shares and average wages for occupations observed in 1979 and comparing the predictions to their actual values in 1979. This procedure generates predictions with a correlation of 0.91 for the true average occupational wage and 0.95 for the true occupational employment share in 1979. I interpret these correlation as strong support for the procedure. I then rank these occupations using their predicted 1979 wages.

The top left panel of Figure 35 confirms the findings of Mishel, Schmitt, and Shierholz (2013) and Lefter and Sand (2011). Job polarization in the 1980s evolved similarly to the 1990s. The main difference comes from job polarization becoming absolute in the 1990s whereas it is only relative in the 1980s. In other words, the middle-ranked or middle-skilled occupations still shrank relative to the lowest ranked occupations even though these lowest ranked occupations contracted relative to all occupations in the 1980s. In the second row of left panel, I show the figure for the period covered by Mishel, Schmitt, and Shierholz (2013). In the third row of left panel, I add the extended years to this figure. In the last row of the left panel, I show changes in employment shares from the dates of occupational coding breaks. Thus, the smoothing the breaks plays no role in shaping this figure. All of these figures suggest a long-run trend towards job polarization. However, the lack of job polarization from 1983 to 1991 suggests this change accelerates in shorter episodes as Hershbein and Kahn (2016) suggests. Figure 36 shows that primarily men drive these patterns across the occupational skill distribution as the patterns become more pronounced if looking at only men.

[137]These occupations compose roughly 10-20% of all occupations from 1983 onwards. Excluding them due to their absence in 1979 is misleading. Figures without them show no relative job polarization in the 1980s.

The right panel of Figure 35 shows corresponding changes in occupational wages. Changes in occupational wages appear to be similar across occupations in the 2000s, polarizing in the 1990s, and expanding in the (early) 1980s. This change parallels wages overall. However, changes in the overall wage distribution will be affected by wage changes within occupational ranks and the concentration of workers across occupations (shown in Figure 37). The latter of which does not appear to change much. In contrast to employment shares, Figure 36 shows that women primarily drive patterns in occupation wages across the occupational skill distribution in the 1980s.

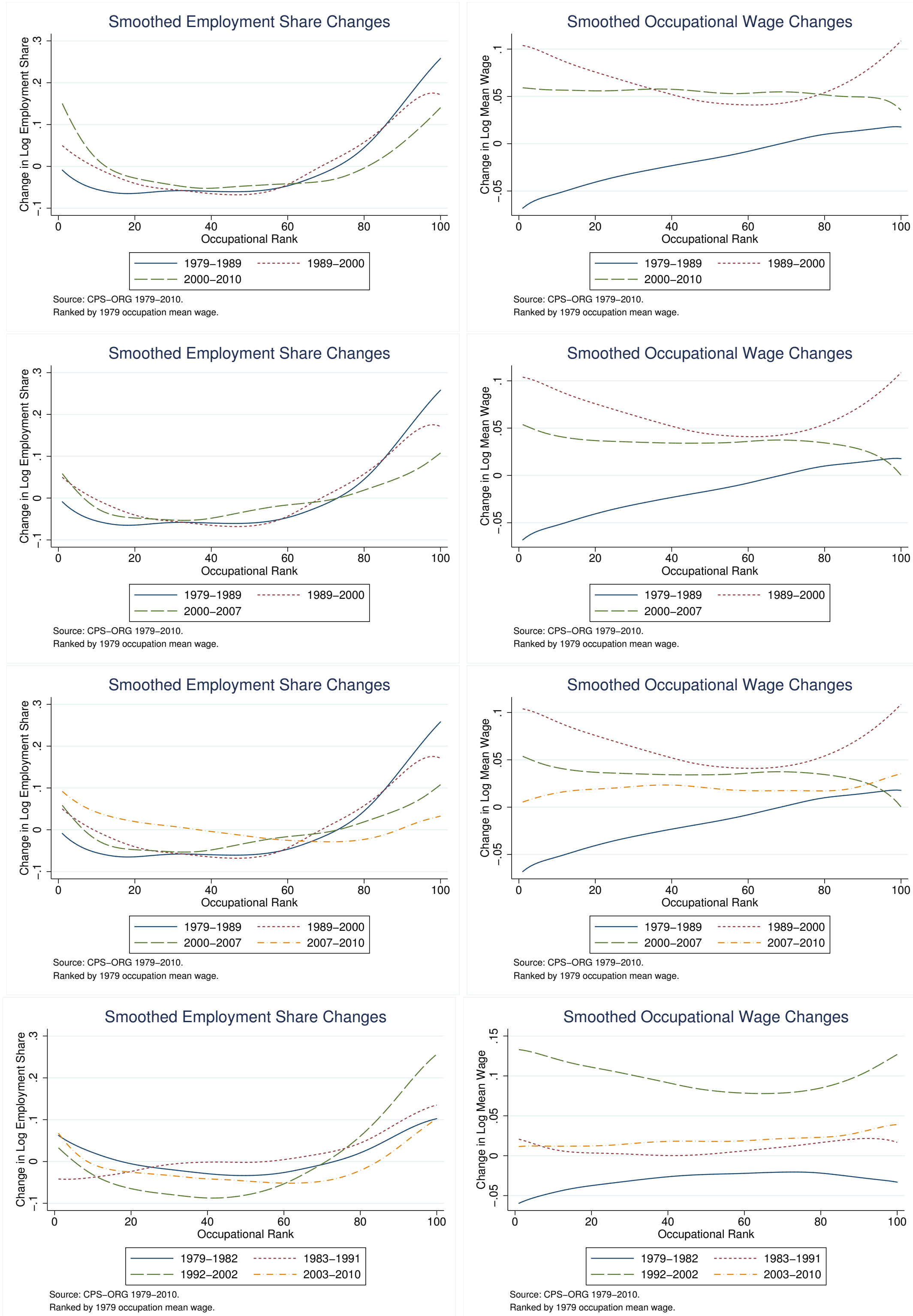


Figure 35: Occupational Employment and Wage Evolution

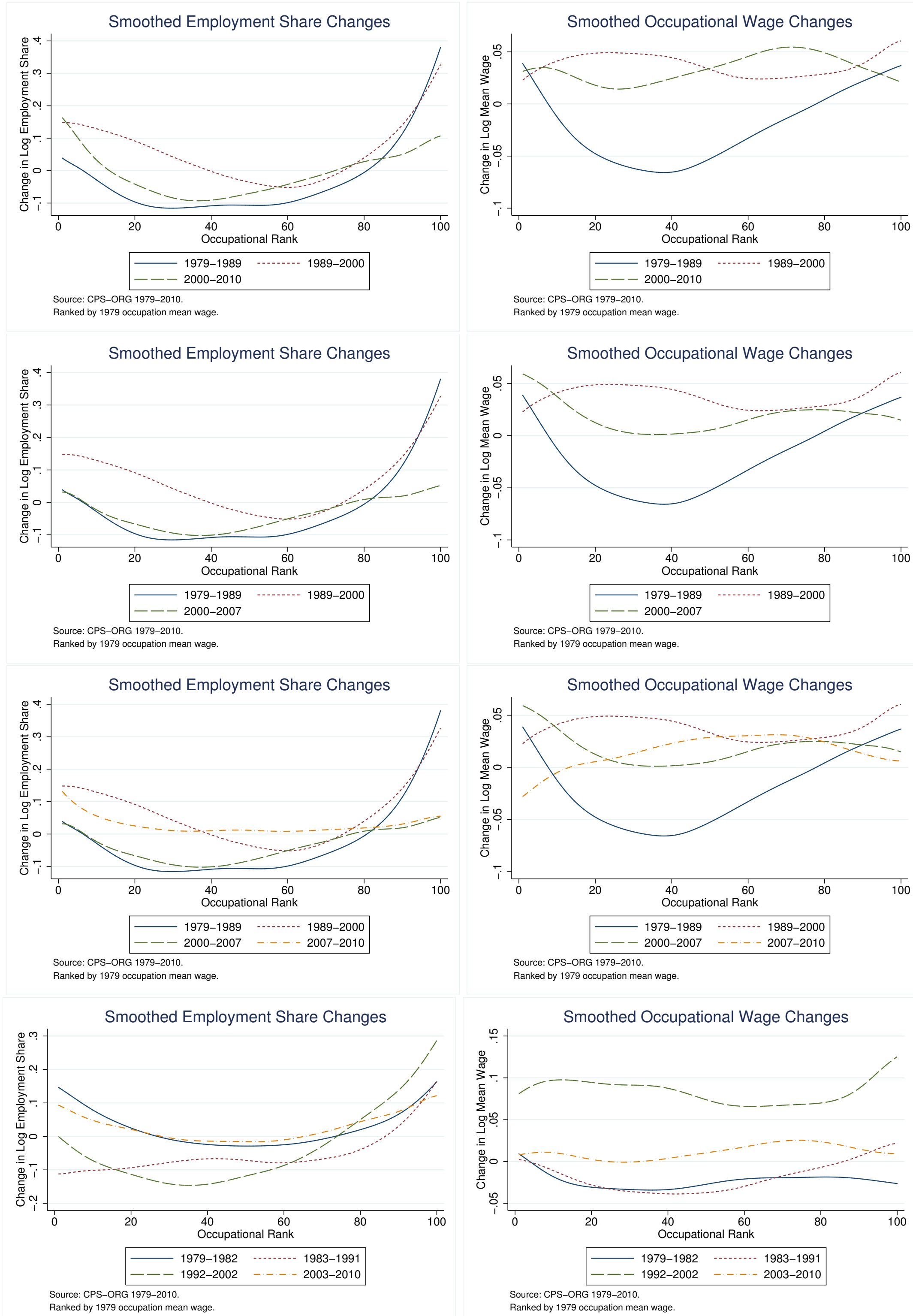


Figure 36: Occupational Employment and Wage Evolution (Men Only)

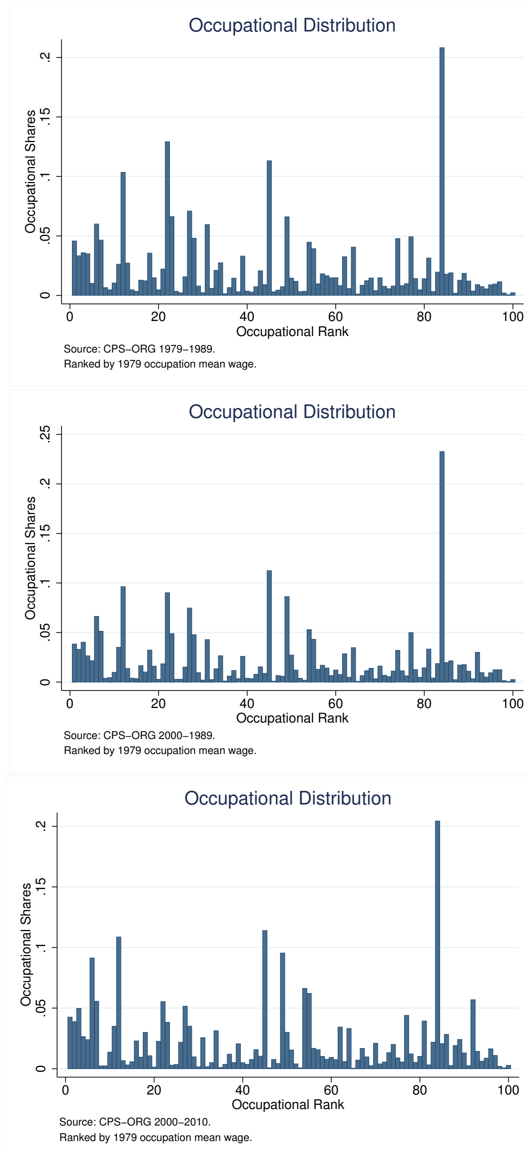


Figure 37: Occupational Distribution across Decades

C Additional Results

Table 17: Model Variants

I	Perfect Foresight Benchmark
II	No Foresight ($\mathbb{E}_t[z_{t+1}] = z_t \ \forall t$)
III	Fixed skill supply distribution ($\mathcal{V}_t(\mathbf{x}) = \mathcal{V}_0(\mathbf{x}) \ \forall t$)
IV	Fixed human capital ($\Gamma_H = 0, \ \Gamma_D = 0$)
V	No matching frictions/Homogeneous specific human capital
VI	$\mathcal{F}(\mathbf{y})$ fixed
VII	$f(\mathbf{x}, \mathbf{y})$ fixed
VIII	$\alpha_M = 0, \ \alpha_{MM} = 0, \ \nu_M = 0, \ \kappa_M = 0$
IX	Nash Bargaining
X	$\mathcal{V}_t(\mathbf{x})$ not adjusted for female labor force participation
XI	Repeated Stationary Model, $\Gamma_D = 0, \ \Gamma_H = 0$

Table 18: Model Accuracy on Targets

	RMSE	Goodness of Fit
I	0.02740	0.946
II	0.02817	0.943
III	0.02532	0.954
IV	0.02370	0.960
V	0.08812	0.445
VI	0.03872	0.893
VII	0.03679	0.903
IX	0.02760	0.946
X	0.02741	0.946

Note: RMSE refers to the root-mean squared error of the model and target moments. Goodness of fit refers to the share of variation in the targets explained by the model.

Table 19: Model Fit (1/2)

	Data	I	II	III
Log Change in Employment Shares				
1979-1989				
High	0.159	0.156	0.162	0.164
Medium	-0.102	-0.098	-0.107	-0.111
Low	0.034	0.014	0.020	0.034
1989-2000				
High	0.171	0.167	0.186	0.170
Medium	-0.125	-0.132	-0.146	-0.135
Low	0.003	0.009	-0.011	0.009
2000-2010				
High	0.026	0.017	0.034	0.021
Medium	-0.039	-0.042	-0.049	-0.041
Low	0.031	0.031	0.025	0.031
Log Change in Occupational Wage				
1979-1989				
High	0.011	0.025	0.016	0.033
Medium	-0.056	-0.023	-0.049	-0.034
Low	-0.078	-0.079	-0.096	-0.080
1989-2000				
High	0.100	0.142	0.090	0.120
Medium	0.050	0.058	0.064	0.052
Low	0.079	0.093	0.115	0.077
2000-2010				
High	0.031	0.057	0.048	0.058
Medium	0.029	0.043	0.030	0.033
Low	-0.029	-0.001	0.010	0.004
Log Change in Wage Percentiles				
1979-1989				
90	0.053	0.034	0.035	0.046
50	-0.018	-0.021	-0.034	-0.032
10	-0.137	-0.127	-0.121	-0.136
1989-2000				
90	0.133	0.130	0.120	0.132
50	0.065	0.112	0.105	0.087
10	0.115	0.114	0.092	0.108
2000-2010				
90	0.091	0.060	0.048	0.065
50	0.026	0.039	0.039	0.029
10	0.011	-0.005	0.002	-0.018

Table 20: Model Fit (2/2)

	Data	I	II	III
Distribution of $\mathbf{y}$				
Mean of y_c				
1980s	0.401	0.405	0.420	0.423
1990s	0.419	0.417	0.449	0.430
2000s	0.432	0.430	0.473	0.442
Standard Deviation of y_c				
1980s	0.204	0.180	0.199	0.193
1990s	0.205	0.182	0.201	0.193
2000s	0.207	0.183	0.203	0.196
Mean of y_m				
1980s	0.436	0.417	0.445	0.432
1990s	0.422	0.403	0.426	0.412
2000s	0.416	0.387	0.409	0.397
Standard Deviation of y_m				
1980s	0.143	0.158	0.144	0.140
1990s	0.146	0.158	0.145	0.141
2000s	0.149	0.159	0.144	0.139
Correlation of (y_c, y_m)				
1980s	-0.031	-0.029	-0.030	-0.022
1990s	-0.079	-0.074	-0.078	-0.080
2000s	-0.114	-0.107	-0.094	-0.112
Log Wage				
Mean				
1980s	2.783	2.782	2.790	2.800
1990s	2.799	2.810	2.822	2.816
2000s	2.896	2.910	2.908	2.891
Standard Deviation				
1980s	0.549	0.578	0.589	0.581
1990s	0.575	0.615	0.620	0.613
2000s	0.598	0.624	0.637	0.629
Distribution of $\mathbf{x}(0)$ and $\mathbf{y}$				
$corr(x_c(0), y_c)$				
1980-1987	0.303	0.403	0.382	0.399
1988-1993	0.457	0.430	0.408	0.419
$corr(x_m(0), y_m)$				
1980-1987	0.078	0.083	0.064	0.065
1988-1993	0.083	0.053	0.050	0.040
Aggregate Job Flows				
Job-to-Job	0.032	0.024	0.035	0.032
Employment-to-Unemployment	0.015	0.016	0.015	0.017
Unemployment-to-Employment	0.261	0.266	0.277	0.262
UE Wage Differential (%)	-0.205	-0.234	-0.273	-0.243
Post-Unemployment Average Wage Drop (%)	-0.264	-0.430	-0.447	-0.417

Table 21: Correlation of Data and Model Wages

	Data	I	II	III	IV	V	VI	VII	VIII	IX	X
Correlation											
All Percentiles											
1979		0.981	0.976	0.978	0.973	0.999	0.976	0.981	0.980	0.985	0.979
1989		0.982	0.980	0.982	0.978	0.998	0.982	0.978	0.981	0.982	0.983
2000		0.964	0.961	0.968	0.958	0.995	0.963	0.950	0.969	0.962	0.966
2010		0.958	0.953	0.961	0.949	0.995	0.954	0.939	0.962	0.957	0.961
Percentiles 5-95											
1979		0.995	0.991	0.993	0.992	0.998	0.993	0.995	0.993	0.997	0.994
1989		0.996	0.994	0.996	0.994	0.999	0.996	0.994	0.994	0.996	0.996
2000		0.989	0.986	0.992	0.986	0.999	0.988	0.983	0.992	0.989	0.990
2010		0.987	0.982	0.989	0.982	0.999	0.985	0.978	0.991	0.987	0.989
Occupational Wages (1979)											
High	25.344	25.532	25.311	25.023	24.774	27.292	24.018	25.211	24.485	24.856	24.951
Medium	18.216	17.715	17.967	17.855	18.393	14.858	18.907	17.335	17.261	17.147	17.990
Low	14.410	15.126	14.411	15.106	14.884	14.072	15.230	15.234	15.962	14.384	14.835

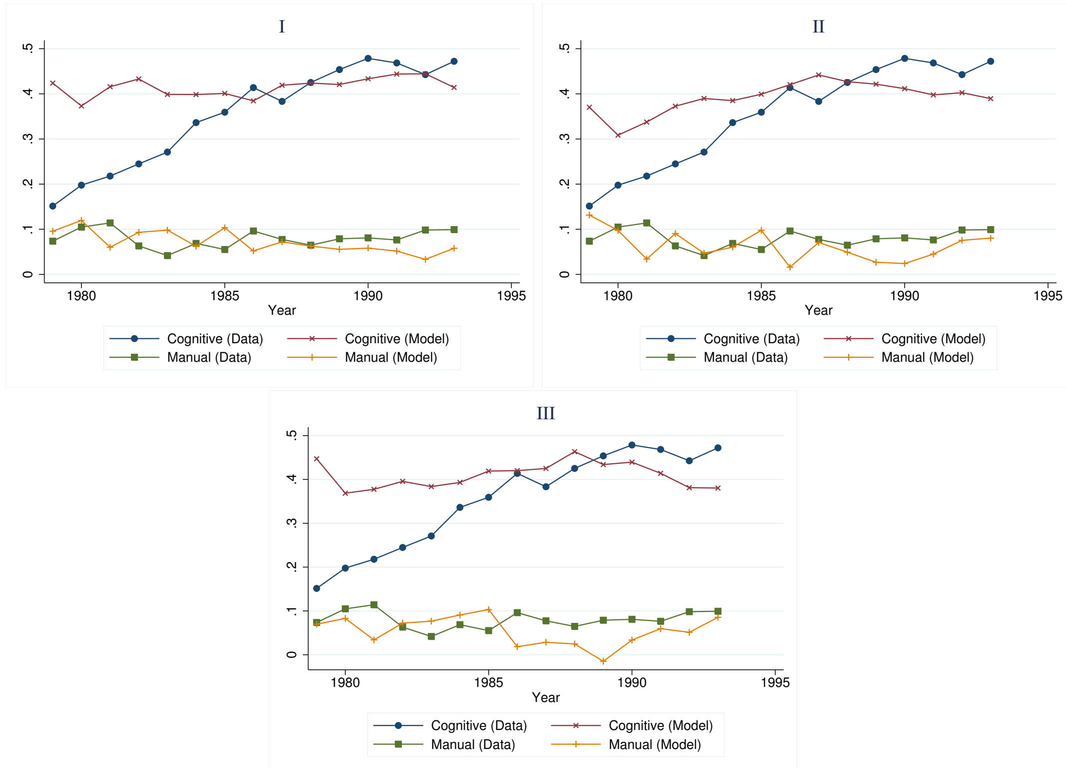


Figure 38: Correlation of $\mathbf{x}(0)$ and $\mathbf{y}$: NLSY79 vs. Model Cohort

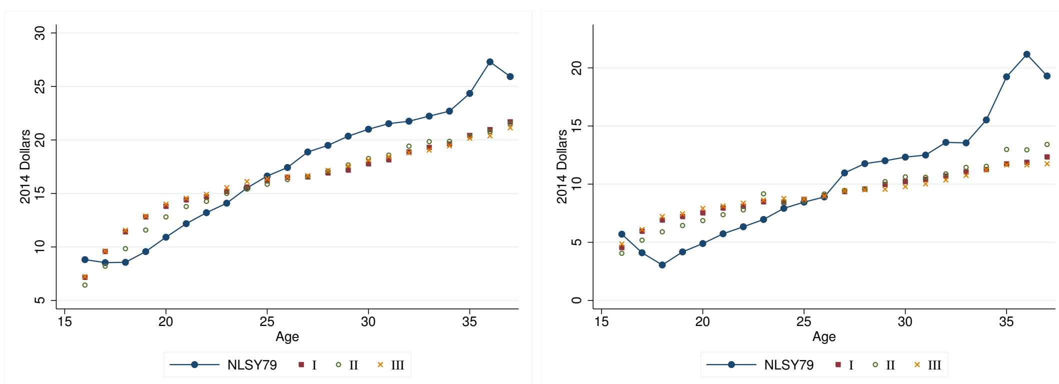


Figure 39: Mean (left) and Standard Deviation (right) Wage-Age Profile

C.1 Demographic Heterogeneity

The model has two demographic dimensions − gender and age. Generally, the model matches aggregate features and changes well but performs less well in capturing the patterns of young workers and different outcomes by gender. For example, the model replicates the pattern of rising and flattening wage dispersion over age (Figure 12). However, it does not match the magnitude of the increase in wage dispersion for young workers compared to the increase seen in the NLSY79 cohort (Appendix Figure 39). The model also replicates average mobility rates and their decline over age but fails to match the initial sharp decline in

Figure 40: Transition Rates by Age

the job-to-job and employment-to-unemployment rates among young workers (Appendix Figure 40).[138] In addition, the model matches the correlation between initial cognitive and manual skills and skill requirements for prime age (30-54) workers. But it fails to capture the increase in this correlation for younger workers as they age (Appendix Figure 38). In essence, young workers appear indistinguishable from prime age workers in terms of endogenous labor market transitions and sorting. Importantly, the employment and wage trends observed remain when restricting to prime age workers in the data, which makes accounting for youth outcomes non-pivotal.

The model distinguishes genders only in the sense that their endowment distributions of cognitive and manual skills differ (Appendix Figure 32). This distinction along these skill dimensions remains insufficient to account for differing occupational employment and wage outcomes between genders as Appendix Figures 41, 42, 43, and 44 show.[139] Only slight differences in emerge for the genders in the model for their pay and allocation of jobs. Whereas the data shows large differences in changes in their occupational wage and employment. For instance, middle-skilled wages rose for women each decade but declined for men in the 1980s. Naturally, only slight differences emerge in the model, because their within-gender marginal distributions of cognitive skills look almost identical. Endowed manual skills differ, but the model judges cognitive skills as far more valuable than manual skills. Furthermore, manual skills adjust rather quickly as

[138] The decline in job-to-job and employment to unemployment switching occurs naturally with ageing, because workers settle into better matches (via human capital accumulation or transitions) as seen in Lise and Postel-Vinay (2016) and Menzio, Telyukova, and Visschers (2016). The decline in unemployment to employment occurs over age, because workers' life expectancy declines. This decline lowers the value of the surplus at any job and increases the chances of exit before finding a job, reesulting in less unemployment to employment transitions with age.

[139] Cortes, Jaimovich, and Siu (2016) document gender differences in terms of exiting the labor market across education levels, which accounts for some of these differences across genders. Workers do not exit based on gender here.

we shall see next. Thus, the gender-education based endowment of skills input into the model fails to result in dramatically different outcomes for men and women.[140]

[140] Another factor like job preferences (Yamaguchi, 2012) or another skill like interpersonal skills (Jaimovich, Siu, and Cortes, 2017) may reconcile differences in occupational wage and employment outcomes for men and women.

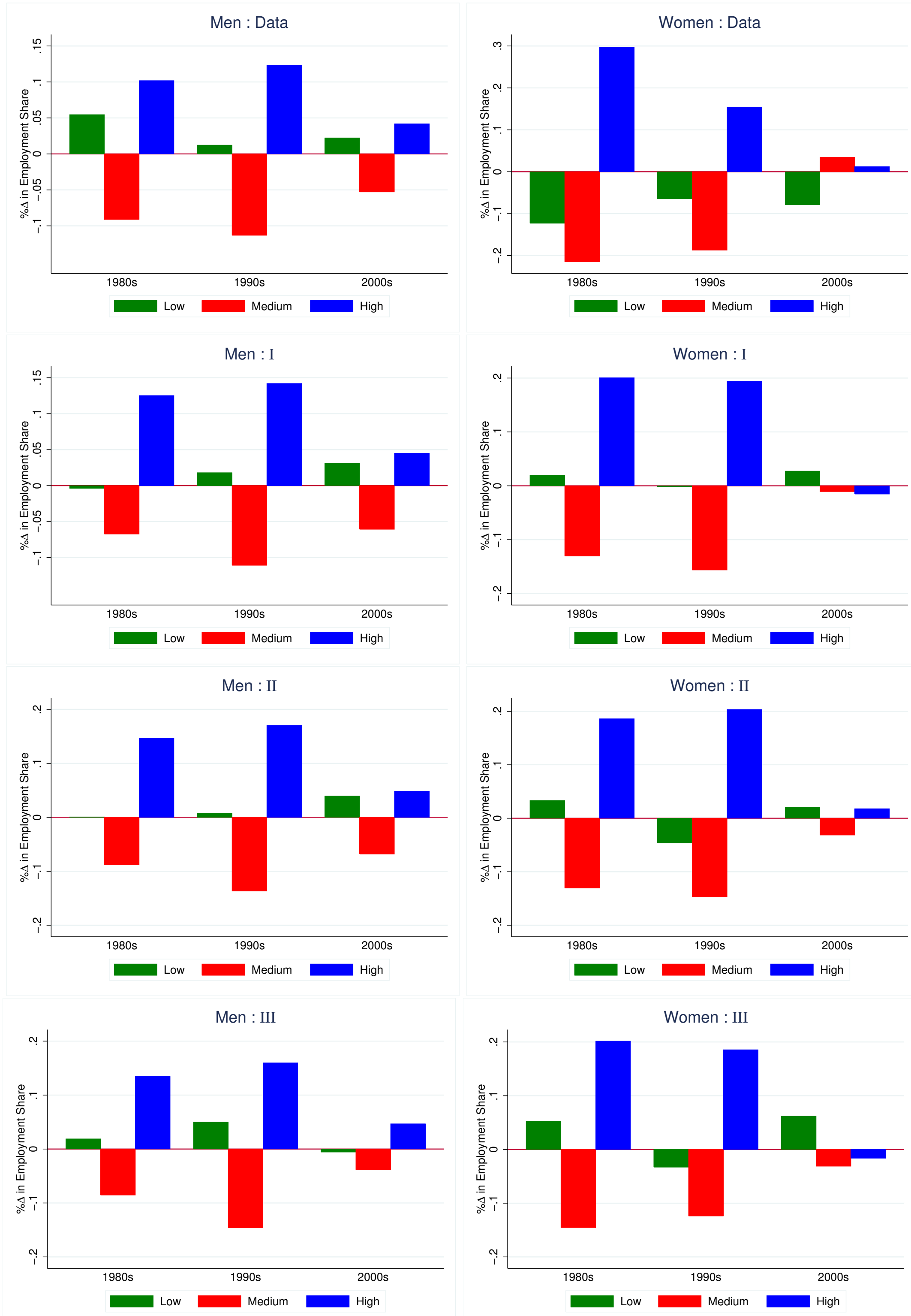


Figure 41: Employment Share Changes: Men (left) vs. Women (right)

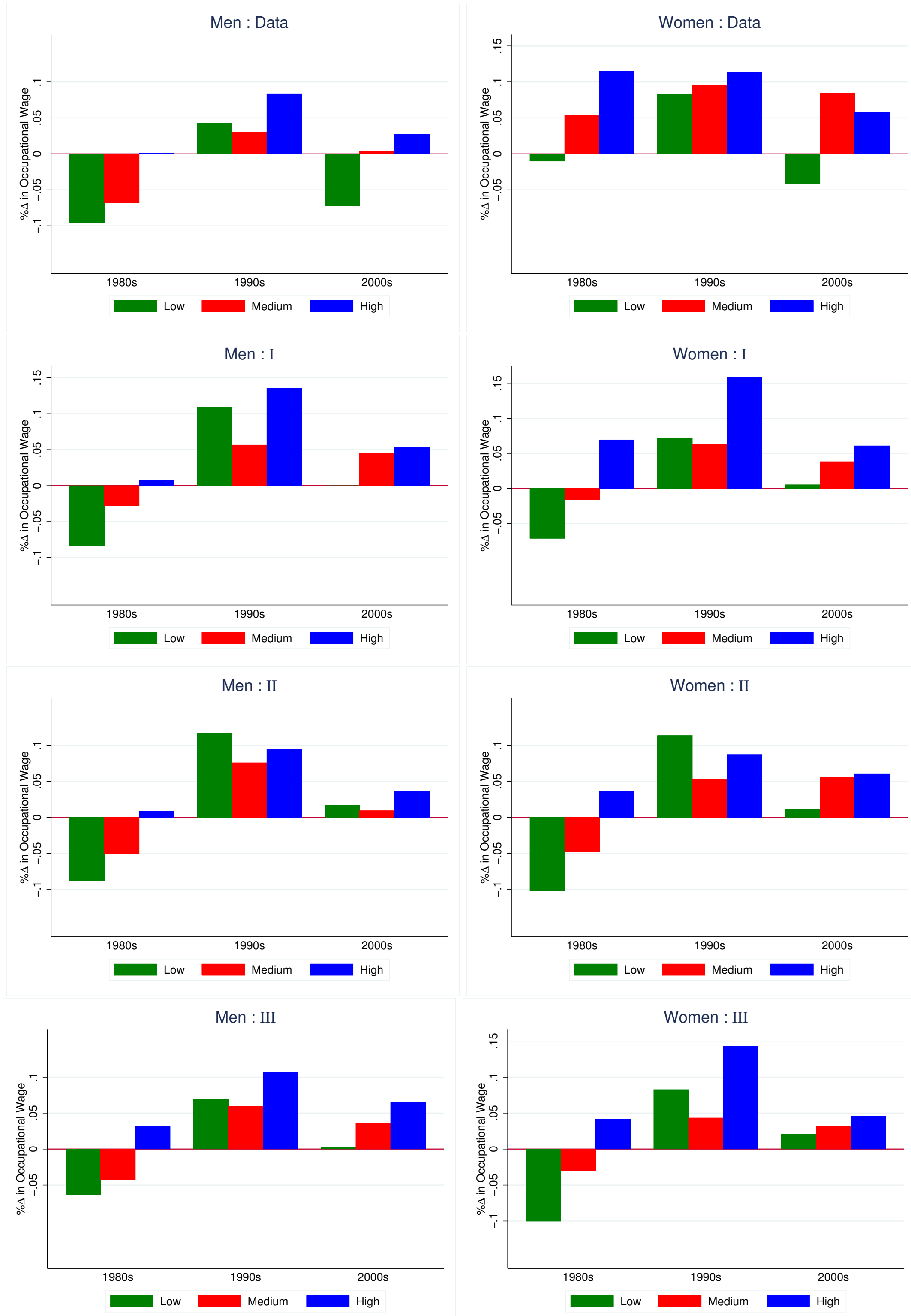


Figure 42: Occupational Wage Changes: Men (left) vs. Women (right)

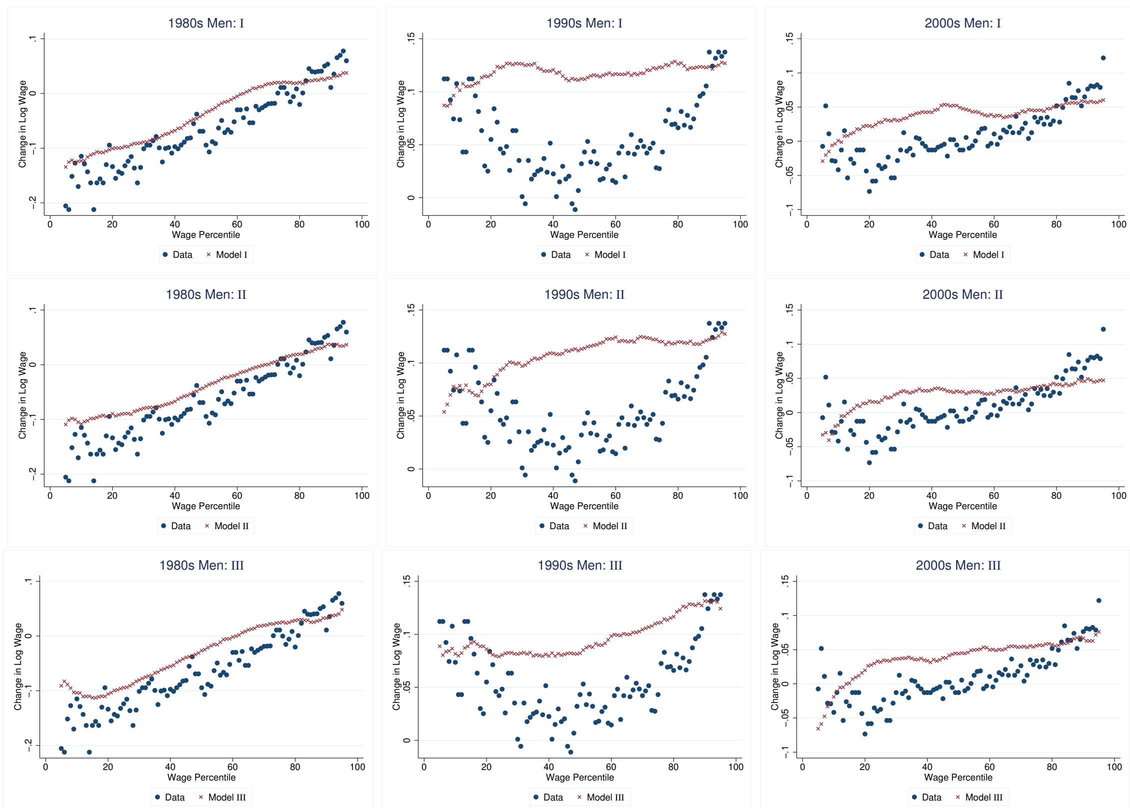


Figure 43: Wage Changes: Men

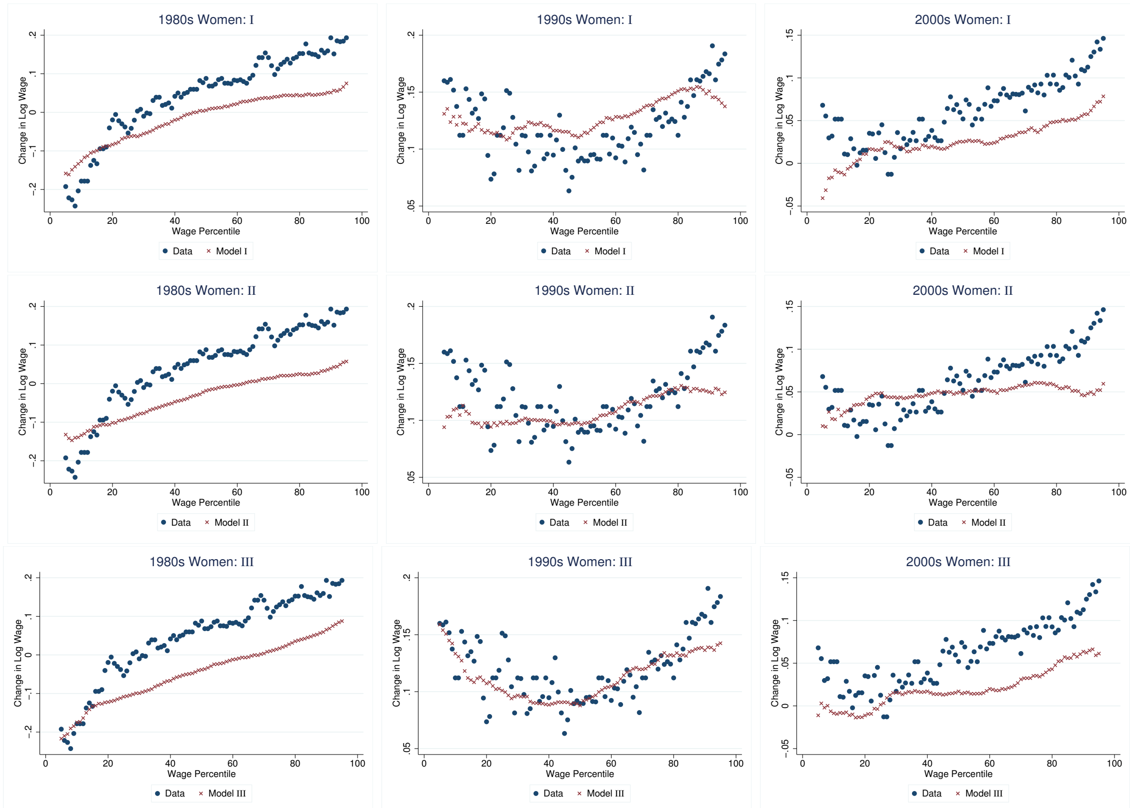


Figure 44: Wage Changes: Women

Figure 45: Density of Log Hourly Wages

Table 22: Time Invariant Parameters

	I	II	III	IV	V	VI	VII	VIII	IX	X
ζ_C	0.837	0.900	0.900	0.896	–	0.900	0.900	0.900	0.837	0.900
ζ_M	0.975	1.100	1.100	1.096	–	1.100	1.100	–	0.975	1.100
b_0	0.683	0.000	0.902	1.812	0.001	1.731	1.302	1.294	1.025	0.888
$\Gamma_H(1,1)$	0.00091	0.00450	0.00286	0	–	0	0	0.0000	0.00091	0.00250
$\Gamma_H(1,2)$	0.00025	0.00200	0.00934	0	–	0	0	0	0.00025	0.01850
$\Gamma_H(2,1)$	0.01850	0.00140	0.01960	0	–	0	0.00180	0	0.02080	0.00612
$\Gamma_H(2,2)$	0.06080	0.05250	0.08970	0	–	0.05060	0.00470	0	0.06050	0.09110
$\Gamma_D(1,1)$	-0.0148	-0.0113	-0.0209	0	–	-0.00550	0	-0.0351	-0.0157	-0.0209
$\Gamma_D(1,2)$	-2.33×10^{-5}	0.0000	-1.45×10^{-5}	0	–	-0.0106	-0.00100	0	-2.33×10^{-5}	-0.00426
$\Gamma_D(2,1)$	-0.0348	-0.0020	-0.0005	0	–	-0.0160	-0.0022	0	-0.0479	-0.00013
$\Gamma_D(2,2)$	-0.0331	-0.0535	-0.0330	0	–	-0.0127	0	0	-0.0290	-0.0309
γ_0	-1.224	-1.161	-1.250	-1.156	-1.340	-1.184	-1.180	-1.095	-1.224	-1.242
γ_1	14.07	14.34	14.06	14.75	16.16	14.93	16.94	14.49	14.04	14.04
γ_2	-13.43	-14.76	-13.62	-15.40	-15.19	-15.68	-14.60	-14.54	-13.16	-13.40
λ	0.425	0.500	0.425	0.419	0.434	0.419	0.416	0.471	0.452	0.425
ν_C	29.71	27.22	38.35	23.87	0	27.33	10.06	13.84	43.66	37.85
ν_M	14.19	0.0004	17.97	0	0	0.0004	20.00	0	20.00	18.91
κ_C	130.8	103.0	128.7	113.0	0	200.0	110.7	100.1	143.8	133.0
κ_M	48.18	47.00	53.62	20.02	0	68.81	73.78	0	48.72	55.81
$\mathbb{M}_u$	0.399	0.380	0.377	0.387	0.270	0.387	0.390	0.339	0.399	0.377
$\mathbb{M}_e$	0.0703	0.0999	0.0751	0.1000	0.0494	0.1000	0.0016	0.0851	0.0703	0.0751
θ_0	0.00289	0.0317	0.0000	0.0377	1.0000	0.0015	0.0003	0.0002	0.00577	0.0000
θ_1	1.000	1.296	1.006	9.978	1.821	1.917	11.290	10.980	1.000	1.006
ω	0.0312	0.0423	0.0469	0.0415	0.0273	0.0319	0.0312	0.0295	0.0312	0.0454

Table 23: $f_t(\mathbf{x}, \mathbf{y})$ Parameters at Sample Dates

	I	II	III	IV	V	VI	VII	VIII	IX	X
$\alpha_{0,t=0}$	1.314	-8×10^{-5}	1.306	0.794	0.176	1.918	-1.970	3.067	1.314	1.306
$\alpha_{0,t=121}$	-1.495	-1.905	-1.479	-1.846	-1.713	-1.082	–	0.744	-1.583	-1.566
$\alpha_{0,t=267}$	-1.950	-1.542	-1.090	-1.985	-3.352	-0.932	–	0.555	-2.040	-1.614
$\alpha_{0,t=335}$	-2.376	-1.348	-1.441	-2.358	-3.505	-1.273	–	0.0373	-2.572	-2.327
$\alpha_{0,t=384}$	-2.683	-1.208	-1.694	-2.627	-3.615	-1.518	–	-0.336	-2.956	-2.841
$\alpha_{C,t=0}$	20.26	19.56	19.23	17.08	1.111	16.11	24.17	17.79	20.26	18.73
$\alpha_{C,t=121}$	20.27	19.77	19.37	17.67	-0.792	14.84	–	18.54	20.27	17.95
$\alpha_{C,t=267}$	19.80	18.18	19.36	17.42	-3.211	14.06	–	18.54	20.51	18.51
$\alpha_{C,t=335}$	19.65	18.16	18.87	17.11	-4.360	13.69	–	18.34	20.73	18.29
$\alpha_{C,t=384}$	19.54	18.14	18.51	16.88	-5.187	13.43	–	18.19	20.90	18.13
$\alpha_{M,t=0}$	-0.775	1.247	-1.283	-0.110	-2.492	0.0360	-0.0169	0	-0.702	-1.283
$\alpha_{M,t=121}$	-0.853	0.646	-1.423	-0.575	0.392	0.0507	–	0	-0.661	-1.491
$\alpha_{M,t=267}$	-0.516	0.571	-1.423	-0.250	7.762	-0.00679	–	0	-1.644	-1.564
$\alpha_{M,t=335}$	-0.0161	0.403	-0.817	0.140	9.815	-0.252	–	0	-1.143	-0.943
$\alpha_{M,t=384}$	0.344	0.282	-0.379	0.421	11.29	-0.429	–	0	-0.783	-0.496
$\alpha_{CC,t=0}$	9.914	10.62	8.379	6.067	25.57	8.373	-2.501	7.444	10.41	9.063
$\alpha_{CC,t=121}$	21.23	16.62	21.01	17.74	35.39	22.77	–	15.76	20.84	22.48
$\alpha_{CC,t=267}$	31.83	24.52	32.68	29.33	41.19	36.96	–	28.82	32.26	33.44
$\alpha_{CC,t=335}$	33.37	26.56	34.95	30.75	43.01	38.21	–	30.58	33.90	35.85
$\alpha_{CC,t=384}$	34.48	28.04	36.58	31.77	44.32	39.12	–	31.85	35.07	37.59
$\alpha_{MM,t=0}$	8.427	8.877	8.615	11.88	-3.117	11.25	7.914	0	8.552	8.631
$\alpha_{MM,t=121}$	9.055	10.14	10.46	12.61	-8.731	12.70	–	0	9.023	11.40
$\alpha_{MM,t=267}$	6.261	6.174	8.193	6.352	-17.44	12.77	–	0	6.228	10.00
$\alpha_{MM,t=335}$	6.069	4.174	7.701	6.305	-19.99	13.81	–	0	6.036	10.24
$\alpha_{MM,t=384}$	5.930	2.733	7.347	6.271	-21.83	14.57	–	0	5.897	10.41

Table 24: $\mathcal{F}_t(\mathbf{y})$ Parameters at Sample Dates

	I	II	III	IV	V	VI	VII	VIII	IX	X
$r_{t=0}$	-0.160	-0.0464	-0.0700	-0.110	-0.264	-0.238	-0.103	-0.0100	-0.169	-0.0700
$r_{t=121}$	-0.240	-0.0974	-0.150	-0.199	-0.405	–	-0.182	0.884	-0.249	-0.135
$r_{t=267}$	-0.313	-0.313	-0.277	-0.303	-0.482	–	-0.199	-0.405	-0.323	-0.342
$r_{t=335}$	-0.236	-0.175	-0.192	-0.142	-0.505	–	-0.199	-0.123	-0.303	-0.193
$r_{t=384}$	-0.181	-0.0753	-0.130	-0.0254	-0.521	–	-0.199	0.0808	-0.290	-0.0850
$a_{C,t=0}$	1.200	1.100	1.200	1.198	1.568	1.200	1.200	1.169	1.231	1.200
$a_{C,t=121}$	1.188	1.100	1.173	1.150	1.556	–	1.285	1.187	1.219	1.113
$a_{C,t=267}$	1.421	1.228	1.405	1.437	1.784	–	1.487	1.331	1.453	1.330
$a_{C,t=335}$	1.387	1.140	1.390	1.385	1.758	–	1.466	1.327	1.418	1.315
$a_{C,t=384}$	1.361	1.076	1.379	1.347	1.738	–	1.451	1.324	1.393	1.304
$b_{C,t=0}$	2.625	2.062	2.156	2.505	4.964	2.500	2.250	1.942	2.656	2.156
$b_{C,t=121}$	2.667	2.062	2.108	2.505	4.960	–	2.194	1.942	2.698	2.123
$b_{C,t=267}$	2.659	2.032	2.062	2.495	5.008	–	2.255	1.940	2.690	2.079
$b_{C,t=335}$	2.614	1.955	1.969	2.388	5.010	–	2.255	1.820	2.646	2.079
$b_{C,t=384}$	2.582	1.900	1.902	2.311	5.012	–	2.255	1.733	2.614	2.079
$a_{M,t=0}$	2.765	3.450	3.242	3.200	3.942	2.820	3.712	3.080	2.765	3.250
$a_{M,t=121}$	2.755	3.306	3.141	3.260	3.924	–	3.579	2.930	2.755	3.164
$a_{M,t=267}$	2.484	3.269	2.873	2.828	3.730	–	3.363	2.758	2.484	2.929
$a_{M,t=335}$	2.454	3.246	2.843	2.773	3.740	–	3.354	2.770	2.454	2.899
$a_{M,t=384}$	2.432	3.229	2.821	2.734	3.746	–	3.347	2.778	2.432	2.877
b_M	6.073	8.987	9.813	8.004	10.87	8.773	14.83	7.913	6.018	9.828

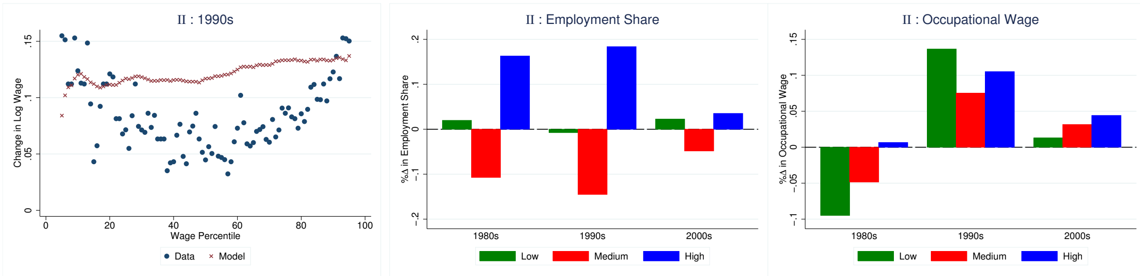


Figure 46: Creating Lower Tail Compression in II

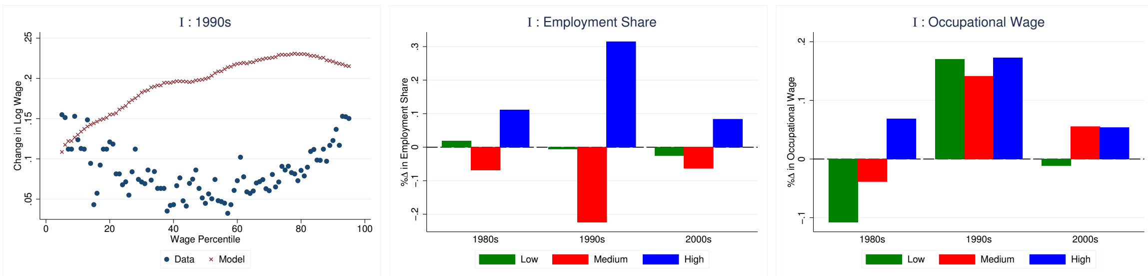


Figure 47: I Estimates with No Foresight

Figure 48: II Estimates with Foresight

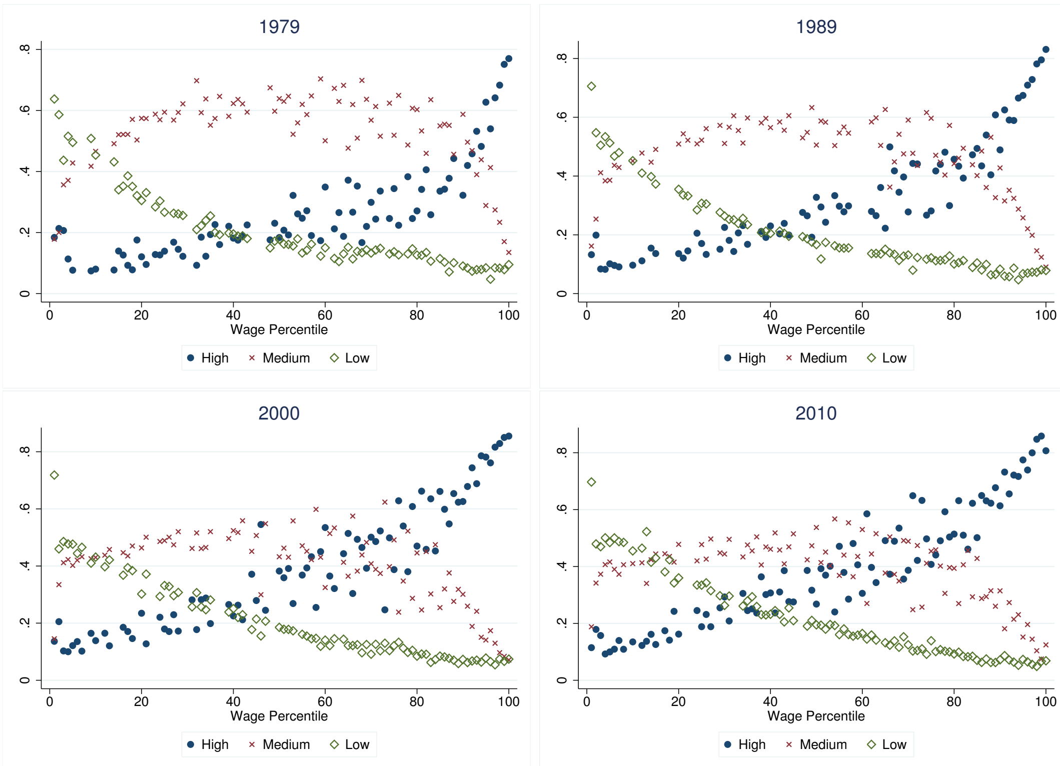


Figure 49: Employment Shares at Wage Percentiles (CPS)

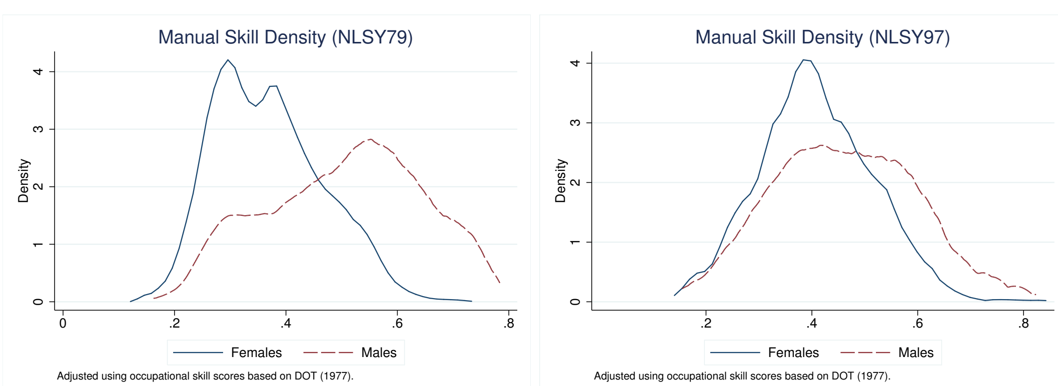


Figure 50: Manual Skill: NLSY79 v. NLSY97

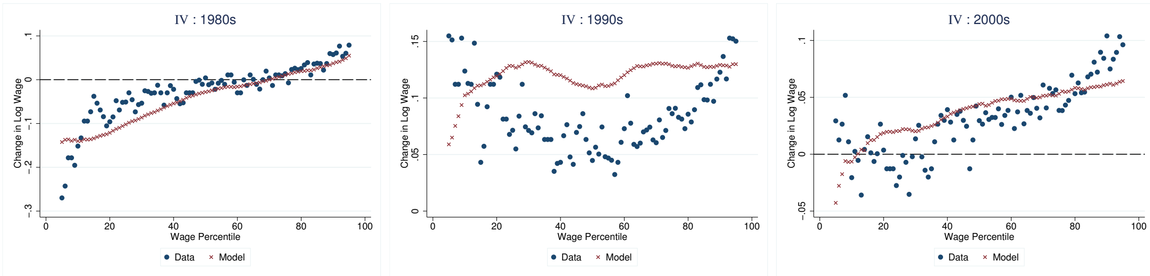


Figure 51: Fixed Specific Human Capital

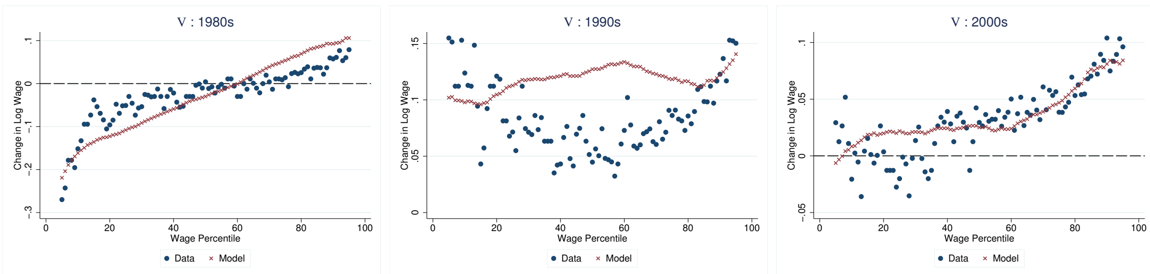


Figure 52: Homogeneous Specific Human Capital

Table 25: Learning Frictions Decomposition (1/2)

	Data	I	IV	V
Log Change in Employment Shares				
1979-1989				
High	0.159	0.156	0.159	0.150
Medium	-0.102	-0.098	-0.094	-0.113
Low	0.034	0.014	0.049	0.028
1989-2000				
High	0.171	0.167	0.181	0.142
Medium	-0.125	-0.132	-0.111	-0.129
Low	0.003	0.009	0.006	0.001
2000-2010				
High	0.026	0.017	0.029	0.029
Medium	-0.039	-0.042	-0.041	-0.044
Low	0.031	0.031	0.037	0.030
Log Change in Occupational Wage				
1979-1989				
High	0.011	0.025	0.042	0.040
Medium	-0.056	-0.023	-0.052	-0.050
Low	-0.078	-0.079	-0.051	-0.102
1989-2000				
High	0.100	0.142	0.136	0.119
Medium	0.050	0.058	0.071	0.033
Low	0.079	0.093	0.083	0.058
2000-2010				
High	0.031	0.057	0.070	0.038
Medium	0.029	0.043	0.036	0.028
Low	-0.029	-0.001	-0.005	-0.031
Log Change in Wage Percentiles				
1979-1989				
90	0.053	0.034	0.037	0.093
50	-0.018	-0.021	-0.028	-0.030
10	-0.137	-0.127	-0.140	-0.162
1989-2000				
90	0.133	0.130	0.129	0.122
50	0.065	0.112	0.108	0.127
10	0.115	0.114	0.102	0.098
2000-2010				
90	0.091	0.060	0.061	0.081
50	0.026	0.039	0.044	0.025
10	0.011	-0.005	-0.007	0.008

Table 26: Learning Frictions Decomposition (2/2)

	Data	I	IV	V
Distribution of $\mathbf{y}$				
Mean of y_C				
1980s	0.401	0.405	0.410	0.496
1990s	0.419	0.417	0.422	0.499
2000s	0.432	0.430	0.437	0.508
Standard Deviation of y_C				
1980s	0.204	0.180	0.183	0.110
1990s	0.205	0.182	0.187	0.112
2000s	0.207	0.183	0.190	0.115
Mean of y_M				
1980s	0.436	0.417	0.449	0.439
1990s	0.422	0.403	0.436	0.420
2000s	0.416	0.387	0.420	0.401
Standard Deviation of y_M				
1980s	0.143	0.158	0.147	0.143
1990s	0.146	0.158	0.148	0.144
2000s	0.149	0.159	0.149	0.142
Correlation of (y_C, y_M)				
1980s	-0.031	-0.029	-0.044	-0.007
1990s	-0.079	-0.074	-0.100	-0.080
2000s	-0.114	-0.107	-0.105	-0.108
Log Wage				
Mean				
1980s	2.783	2.782	2.792	2.787
1990s	2.799	2.810	2.831	2.810
2000s	2.896	2.910	2.914	2.892
Standard Deviation				
1980s	0.549	0.578	0.577	0.504
1990s	0.575	0.615	0.615	0.565
2000s	0.598	0.624	0.632	0.589
Distribution of $\mathbf{x}(0)$ and $\mathbf{y}$				
$corr(x_C(0), y_C)$				
1980-1987	0.303	0.403	0.394	.
1988-1993	0.457	0.430	0.446	.
$corr(x_M(0), y_M)$				
1980-1987	0.078	0.083	0.094	.
1988-1993	0.083	0.053	0.095	.
Aggregate Job Flows				
Job-to-Job	0.030	0.019	0.030	0.009
Employment-to-Unemployment	0.015	0.016	0.016	0.021
Unemployment-to-Employment	0.261	0.266	0.270	0.104
U-to-E Wage Differential (%)	-0.205	-0.234	-0.243	-0.152
Unemployment Spell Average Wage Drop (%)	-0.264	-0.430	-0.381	-0.504

Table 27: $\mathcal{F}_t(\mathbf{y})$ and $f_t(\mathbf{x},\mathbf{y})$ Decomposition (1/2)

	Data	I	VI	VII
Log Change in Employment Shares				
1979-1989				
High	0.159	0.156	0.157	0.167
Medium	-0.102	-0.098	-0.037	-0.094
Low	0.034	0.014	-0.060	-0.005
1989-2000				
High	0.171	0.167	0.075	0.158
Medium	-0.125	-0.132	-0.046	-0.111
Low	0.003	0.009	-0.006	-0.013
2000-2010				
High	0.026	0.017	-0.053	0.044
Medium	-0.039	-0.042	0.021	-0.038
Low	0.031	0.031	0.016	-0.002
Log Change in Occupational Wage				
1979-1989				
High	0.011	0.025	0.045	0.028
Medium	-0.056	-0.023	-0.016	0.018
Low	-0.078	-0.079	-0.080	0.006
1989-2000				
High	0.100	0.142	0.119	0.037
Medium	0.050	0.058	0.099	0.086
Low	0.079	0.093	0.082	0.088
2000-2010				
High	0.031	0.057	0.059	0.032
Medium	0.029	0.043	0.010	0.028
Low	-0.029	-0.001	0.038	0.030
Log Change in Wage Percentiles				
1979-1989				
90	0.053	0.034	0.054	0.049
50	-0.018	-0.021	-0.027	0.044
10	-0.137	-0.127	-0.137	0.015
1989-2000				
90	0.133	0.130	0.088	0.067
50	0.065	0.112	0.107	0.097
10	0.115	0.114	0.154	0.121
2000-2010				
90	0.091	0.060	0.043	0.031
50	0.026	0.039	0.025	0.036
10	0.011	-0.005	0.002	0.063

Table 28: $\mathcal{F}_t(\mathbf{y})$ and $f_t(\mathbf{x}, \mathbf{y})$ Decomposition (2/2)

	Data	I	VI	VII
Distribution of $\mathbf{y}$				
Mean of y_C				
1980s	0.401	0.405	0.383	0.404
1990s	0.419	0.417	0.392	0.426
2000s	0.432	0.430	0.391	0.444
Standard Deviation of y_C				
1980s	0.204	0.180	0.176	0.192
1990s	0.205	0.182	0.176	0.194
2000s	0.207	0.183	0.176	0.194
Mean of y_M				
1980s	0.436	0.417	0.390	0.421
1990s	0.422	0.403	0.388	0.409
2000s	0.416	0.387	0.388	0.398
Standard Deviation of y_M				
1980s	0.143	0.158	0.141	0.121
1990s	0.146	0.158	0.142	0.122
2000s	0.149	0.159	0.141	0.122
Correlation of (y_C, y_M)				
1980s	-0.031	-0.029	-0.066	-0.049
1990s	-0.079	-0.074	-0.057	-0.085
2000s	-0.114	-0.107	-0.071	-0.118
Log Wage				
Mean				
1980s	2.783	2.782	2.783	2.796
1990s	2.799	2.810	2.823	2.859
2000s	2.896	2.910	2.897	2.933
Standard Deviation				
1980s	0.549	0.578	0.567	0.596
1990s	0.575	0.615	0.583	0.589
2000s	0.598	0.624	0.582	0.578
Distribution of $\mathbf{x}(0)$ and $\mathbf{y}$				
$corr(x_C(0), y_C)$				
1980-1987	0.303	0.403	0.447	0.400
1988-1993	0.457	0.430	0.475	0.424
$corr(x_M(0), y_M)$				
1980-1987	0.078	0.083	0.107	0.164
1988-1993	0.083	0.053	0.097	0.123
Aggregate Job Flows				
Job-to-Job	0.030	0.019	0.026	0.001
Employment-to-Unemployment	0.015	0.016	0.016	0.017
Unemployment-to-Employment	0.261	0.266	0.253	0.251
U-to-E Wage Differential (%)	-0.205	-0.234	-0.267	-0.185
Unemployment Spell Average Wage Drop (%)	-0.264	-0.430	-0.401	-0.393

Table 29: Skil Content Decomposition (1/2)

	Data	I	III	X	VIII
Log Change in Employment Shares					
1979-1989					
High	0.159	0.156	0.164	0.151	0.189
Medium	-0.102	-0.098	-0.112	-0.102	-0.107
Low	0.034	0.014	0.035	0.038	0.021
1989-2000					
High	0.171	0.167	0.171	0.165	0.148
Medium	-0.125	-0.132	-0.135	-0.128	-0.120
Low	0.003	0.009	0.009	0.016	0.033
2000-2010					
High	0.026	0.017	0.022	0.021	0.036
Medium	-0.039	-0.042	-0.040	-0.036	-0.044
Low	0.031	0.031	0.030	0.027	0.023
Log Change in Occupational Wage					
1979-1989					
High	0.011	0.025	0.032	0.032	0.051
Medium	-0.056	-0.023	-0.035	-0.015	-0.014
Low	-0.078	-0.079	-0.081	-0.079	-0.111
1989-2000					
High	0.100	0.142	0.121	0.135	0.144
Medium	0.050	0.058	0.052	0.066	0.038
Low	0.079	0.093	0.077	0.090	0.137
2000-2010					
High	0.031	0.057	0.059	0.053	0.091
Medium	0.029	0.043	0.036	0.044	0.055
Low	-0.029	-0.001	0.006	0.016	-0.023
Log Change in Wage Percentiles					
1979-1989					
90	0.053	0.034	0.045	0.041	0.053
50	-0.018	-0.021	-0.033	-0.017	-0.008
10	-0.137	-0.127	-0.137	-0.099	-0.154
1989-2000					
90	0.133	0.130	0.132	0.133	0.152
50	0.065	0.112	0.087	0.099	0.087
10	0.115	0.114	0.107	0.116	0.105
2000-2010					
90	0.091	0.060	0.066	0.062	0.095
50	0.026	0.039	0.030	0.039	0.045
10	0.011	-0.005	-0.015	0.005	-0.001

Table 30: Skill Content Decomposition (2/2)

	Data	I	III	X	VIII
Distribution of $\mathbf{y}$					
Mean of y_C					
1980s	0.401	0.405	0.423	0.419	0.419
1990s	0.419	0.417	0.430	0.426	0.426
2000s	0.432	0.430	0.442	0.439	0.429
Standard Deviation of y_C					
1980s	0.204	0.180	0.193	0.191	0.196
1990s	0.205	0.182	0.193	0.193	0.198
2000s	0.207	0.183	0.196	0.194	0.199
Mean of y_M					
1980s	0.436	0.417	0.432	0.437	0.449
1990s	0.422	0.403	0.413	0.416	0.431
2000s	0.416	0.387	0.397	0.401	0.405
Standard Deviation of y_M					
1980s	0.143	0.158	0.140	0.140	0.144
1990s	0.146	0.158	0.141	0.140	0.146
2000s	0.149	0.159	0.139	0.140	0.149
Correlation of (y_C, y_M)					
1980s	-0.031	-0.029	-0.022	-0.018	0.141
1990s	-0.079	-0.074	-0.080	-0.080	0.133
2000s	-0.114	-0.107	-0.112	-0.117	-0.138
Log Wage					
Mean					
1980s	2.783	2.782	2.799	2.797	2.781
1990s	2.799	2.810	2.815	2.829	2.827
2000s	2.896	2.910	2.890	2.912	2.905
Standard Deviation					
1980s	0.549	0.578	0.581	0.579	0.573
1990s	0.575	0.615	0.613	0.605	0.610
2000s	0.598	0.624	0.630	0.620	0.633
Distribution of $\mathbf{x}(0)$ and $\mathbf{y}$					
$corr(x_C(0), y_C)$					
1980-1987	0.303	0.403	0.398	0.407	0.371
1988-1993	0.457	0.430	0.419	0.415	0.406
$corr(x_M(0), y_M)$					
1980-1987	0.078	0.083	0.063	0.074	0.026
1988-1993	0.083	0.053	0.040	0.037	0.030
Aggregate Job Flows					
Job-to-Job	0.030	0.019	0.021	0.020	0.027
Employment-to-Unemployment	0.015	0.016	0.017	0.017	0.014
Unemployment-to-Employment	0.261	0.266	0.262	0.256	0.267
U-to-E Wage Differential (%)	-0.205	-0.273	-0.243	-0.247	-0.291
Unemployment Spell Average Wage Drop (%)	-0.264	-0.447	-0.417	-0.430	-0.417

Table 31: Nash Bargaining (1/2)

	Data	I	IX
Log Change in Employment Shares			
1979-1989			
High	0.159	0.156	0.154
Medium	-0.102	-0.098	-0.108
Low	0.034	0.014	0.026
1989-2000			
High	0.171	0.167	0.161
Medium	-0.125	-0.132	-0.117
Low	0.003	0.009	-0.011
2000-2010			
High	0.026	0.017	0.020
Medium	-0.039	-0.042	-0.046
Low	0.031	0.031	0.033
Log Change in Occupational Wage			
1979-1989			
High	0.011	0.025	0.019
Medium	-0.056	-0.023	-0.018
Low	-0.078	-0.079	-0.059
1989-2000			
High	0.100	0.142	0.120
Medium	0.050	0.058	0.046
Low	0.079	0.093	0.093
2000-2010			
High	0.031	0.057	0.057
Medium	0.029	0.043	0.030
Low	-0.029	-0.001	0.013
Log Change in Wage Percentiles			
1979-1989			
90	0.053	0.034	0.040
50	-0.018	-0.021	-0.007
10	-0.137	-0.127	-0.127
1989-2000			
90	0.133	0.130	0.112
50	0.065	0.112	0.107
10	0.115	0.114	0.105
2000-2010			
90	0.091	0.060	0.057
50	0.026	0.039	0.039
10	0.011	-0.005	0.001

Table 32: Nash Bargaining (2/2)

	Data	I	IX
Distribution of $\mathbf{y}$			
Mean of y_C			
1980s	0.401	0.405	0.411
1990s	0.419	0.417	0.422
2000s	0.432	0.430	0.435
Standard Deviation of y_C			
1980s	0.204	0.180	0.176
1990s	0.205	0.182	0.178
2000s	0.207	0.183	0.180
Mean of y_M			
1980s	0.436	0.417	0.415
1990s	0.422	0.403	0.402
2000s	0.416	0.387	0.387
Standard Deviation of y_M			
1980s	0.143	0.158	0.158
1990s	0.146	0.158	0.157
2000s	0.149	0.159	0.158
Correlation of (y_C, y_M)			
1980s	-0.031	-0.029	-0.014
1990s	-0.079	-0.074	-0.066
2000s	-0.114	-0.107	-0.129
Log Wage			
Mean			
1980s	2.783	2.782	2.760
1990s	2.799	2.810	2.782
2000s	2.896	2.910	2.882
Standard Deviation			
1980s	0.549	0.578	0.567
1990s	0.575	0.615	0.603
2000s	0.598	0.624	0.605
Distribution of $\mathbf{x}(0)$ and $\mathbf{y}$			
$corr(x_C(0), y_C)$			
1980-1987	0.303	0.403	0.426
1988-1993	0.457	0.430	0.442
$corr(x_M(0), y_M)$			
1980-1987	0.078	0.083	0.089
1988-1993	0.083	0.053	0.054
Aggregate Job Flows			
Job-to-Job	0.030	0.019	0.018
Employment-to-Unemployment	0.015	0.016	0.016
Unemployment-to-Employment	0.261	0.266	0.254
U-to-E Wage Differential (%)	-0.205	-0.273	-0.234
Unemployment Spell Average Wage Drop (%)	-0.264	-0.447	-0.431

Figure 53: Nash Bargaining

Table 33: Repeated Stationary Model (1/3)

	Data	XI
Log Change in Employment Shares		
1979-1989		
High	0.159	0.167
Medium	-0.102	-0.082
Low	0.034	-0.021
1989-2000		
High	0.171	0.202
Medium	-0.125	-0.159
Low	0.003	0.011
2000-2010		
High	0.026	-0.082
Medium	-0.039	0.081
Low	0.031	-0.021
Log Change in Occupational Wage		
1979-1989		
High	0.011	0.031
Medium	-0.056	-0.105
Low	-0.078	-0.041
1989-2000		
High	0.100	0.053
Medium	0.050	0.093
Low	0.079	0.078
2000-2010		
High	0.031	0.045
Medium	0.029	0.022
Low	-0.029	0.031
Log Change in Wage Percentiles		
1979-1989		
90	0.053	0.005
50	-0.018	-0.026
10	-0.137	-0.214
1989-2000		
90	0.133	0.153
50	0.065	0.062
10	0.115	0.128
2000-2010		
90	0.091	0.016
50	0.026	0.022
10	0.011	-0.027

Table 34: Repeated Stationary Model (2/3)

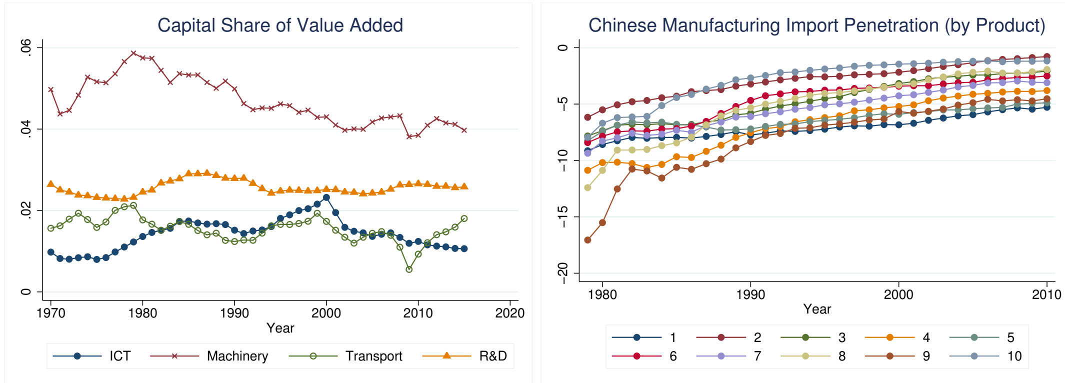
	Data	XI
Distribution of $\mathbf{y}$		
Mean of y_C		
1979	0.388	0.402
1989	0.411	0.419
2000	0.426	0.439
2010	0.436	0.436
Standard Deviation of y_C		
1979	0.206	0.189
1989	0.203	0.197
2000	0.206	0.208
2010	0.208	0.208
Mean of y_M		
1979	0.445	0.447
1989	0.428	0.441
2000	0.418	0.417
2010	0.413	0.437
Standard Deviation of y_M		
1979	0.143	0.139
1989	0.144	0.151
2000	0.148	0.138
2010	0.150	0.164
Correlation of (y_C, y_M)		
1979	-0.017	0.017
1989	-0.068	-0.068
2000	-0.118	-0.118
2010	-0.111	-0.017
Log Wage		
Mean		
1979	2.810	2.827
1989	2.783	2.763
2000	2.872	2.859
2010	2.906	2.876
Standard Deviation		
1979	0.511	0.515
1989	0.572	0.609
2000	0.583	0.609
2010	0.613	0.620

Table 35: Repeated Stationary Model (3/3)

	Data	XI
Aggregate Job Flows		
Job-to-Job		
1979	0.030	0.022
1989	0.030	0.027
2000	0.030	0.020
2010	0.030	0.019
Employment-to-Unemployment		
1979	0.015	0.015
1989	0.014	0.012
2000	0.011	0.012
2010	0.016	0.019
Unemployment-to-Employment		
1979	0.291	0.258
1989	0.299	0.290
2000	0.323	0.365
2010	0.162	0.166

Table 36: Repeated Stationary Model Parameters

	1979	1989	2000	2010
ζ_C	0.787	0.892	0.900	0.892
ζ_M	1.100	1.100	1.100	1.100
b_0	2.544	0.002	2.419	2.456
λ	0.425	0.312	0.412	0.482
γ_0	-1.165	-1.339	-1.153	-1.224
γ_1	14.671	15.014	14.917	15.014
γ_2	-15.527	-16.331	-15.983	-16.194
α_0	0.530	1.059	2.303	0.383
α_C	14.266	15.781	3.500	11.220
α_M	0.049	-0.332	1.222	0.287
α_{CC}	12.055	22.512	28.714	32.897
α_{MM}	12.758	7.435	6.980	5.155
ν_C	25.529	39.409	28.561	27.645
ν_M	20.000	0.000	5.888	19.751
κ_C	92.232	109.517	124.606	149.196
κ_M	59.777	38.885	92.336	71.590
$\mathbb{M}_u$	0.600	0.417	0.867	0.482
$\mathbb{M}_e$	0.120	0.120	0.120	0.120
r	-0.285	-0.317	-0.500	-0.450
a_C	0.800	1.102	0.700	0.500
b_C	2.400	2.337	1.700	1.550
a_M	3.500	3.071	3.400	2.951
b_M	7.700	6.000	6.984	4.000
θ_0	0.044	0.001	0.087	0.062
θ_1	2.035	1.008	4.516	14.610
ω	0.010	0.000	0.000	0.019



Note: Manufacturing product groups are 1) food and tobacco, 2) textiles and appliances, 3) wood and furniture, 4) paper and printing, 5) chemicals and petroleum, 6) clay, stone, rubber and leather, 7) metals, 8) equipment, 9) transport, and 10) other products (e.g. toys). Capital share sectors are 1) agriculture, forestry, fishing, and hunting, 2) mining, 3) construction, 4) manufacturing, 5) wholesale and retail trade, 6) transportation and utilities, 7) information and communications, 8) financial, professional and business services, 9) educational and health services, 10) leisure and hospitality, and 11) other services.

Figure 54: Explantory Factors for $\Delta\mathcal{F}(\mathbf{y})$

Table 37: Average Task Content by Occupational Group (1979)

	High	Medium	Low
Offshoring Vulnerability	0.425	-0.310	0.148
Routine Intensity	-1.246	0.900	-0.025
Interpersonal Intensity	0.863	-0.613	-0.678

Table 38: Correlation in Task Content (1979)

	Offshoring Vulnerability	Routine Intensity
Routine Intensity	-0.200	
Interpersonal Intensity	-0.060	-0.608

C.2 $\Delta\mathcal{F}_t(\mathbf{y})$ vs. Δ in Equilibrium Distribution of $\mathbf{y}$

The distribution of skill demand, $\mathcal{F}_t(\mathbf{y})$, serves as the object of interest to infer why skill demand changed here, because the distribution of $\mathbf{y}$ may not reflect skill demand changes. Most reduced-form studies infer demand changes from the equilibrium wage and employment share changes. If $\mathcal{F}_t(\mathbf{y})$ governs the equilibrium distribution of $\mathbf{y}$, then why estimate at $\mathcal{F}_t(\mathbf{y})$? After all, the object remains difficult to estimate and the equilibrium distribution of $\mathbf{y}$ is available with some caveats. However, selection effects (or sorting) in equilibrium lead to changes in the equilibrium distribution of $\mathbf{y}$. Also, skill mismatch, changes in the distribution of $\mathbf{x}$, and search frictions all affect the equilibrium distribution of $\mathbf{y}$. Hence, the observed equilibrium distribution does not necessarily reflect concurrent skill demand everywhere. Figure 55 shows

Table 39: Average Industry Concentration by Occupational Group (1979)

	High	Medium	Low
Agriculture, Forestry, Fishing, & Hunting	0.001	0.004	0.003
Mining	0.009	0.013	0.003
Construction	0.016	0.096	0.006
Manufacturing	0.151	0.351	0.161
Wholesale & Retail Trade	0.047	0.113	0.379
Transportation & Utilities	0.027	0.071	0.070
Information Services	0.014	0.021	0.006
Financial, Professional, & Business Services	0.230	0.086	0.068
Education and Health Services	0.427	0.156	0.144
Leisure & Hospitality	0.005	0.006	0.006
Other Services	0.005	0.030	0.130

Table 40: 1979 Task Content Variance Decomposition on $\Delta\mathcal{F}(\mathbf{y})$

	I	II	III
Offshoring Vulnerability	0.051	0.058	0.128
Routine Intensity	0.025	0.007	0.023
Interpersonal Intensity	0.139	0.070	0.239
Total Variance Contribution	21.9%	10.9%	33.5%

contour plots of the change in the distribution of equilibrium skill requirements and $\mathcal{F}_t(\mathbf{y})$ for the model (III). The model equilibrium distribution of $\mathbf{y}$ appears rather misleading compared to $\mathcal{F}_t(\mathbf{y})$. Skill demands in the model polarize much more than the equilibrium distribution of $\mathbf{y}$ suggests. This difference illustrates why we must look at $\mathcal{F}_t(\mathbf{y})$ directly to judge how skill demands evolved.

The data's equilibrium distribution of $\mathbf{y}$ exhibits polarization although not as strong as the model's skill demands suggests (Figure 56).[141] One interpretation of this difference is the model overestimates the importance of frictions and selection effects, making the equilibrium distribution of $\mathbf{y}$ an imperfect but suitable proxy for $\mathcal{F}_t(\mathbf{y})$. Another interpretation of this difference comes from the construction of $\mathbf{y}$ in the data versus the model. $\mathbf{y}$ changes little within occupations in the data over time, because the DOT waves only took place in 1977 and 1991. We also do not observe dispersion in $\mathbf{y}$ within occupations due to its construction at the occupational level. This aggregation means any change in the area will occur roughly in the same place in the data, whereas changes in an area can be more spread out in the model. This spreading out within occupations makes skill demand polarization more difficult to see.[142] On one hand, aggregation causes the data to better reflects polarizing in skill demands. On the other, it reduces our power to distill between various theories as well as demand shifts and selection effects, lessening the credibility of inference directly from $\mathbf{y}$ in the data.

[141] The model only matches the first and second moments of this distribution.

[142] Collapsing the model's equilibrium distribution of $\mathbf{y}$ into (y_C, y_M) cells shows more polarization. Hence, Figure 56 cannot rule out the possiblity of a strong role for frictions and selection given the strong possibility that data construction drives it.

Table 41: Capital Input and Imports Variance Decomposition on $\Delta\mathcal{F}(\mathbf{y})$

	I	II	III
Δ Chinese Manufacturing Import Penetration	0.000	0.014	0.003
Δ Capital Investment			
Information & Communications Technology	0.001	0.002	0.006
Machinery	0.026	0.023	0.039
Research & Development	0.076	0.047	0.094
Transportation Equipment	0.050	0.001	0.104
Total Variance Contribution	58.8%	28.4%	56.9%

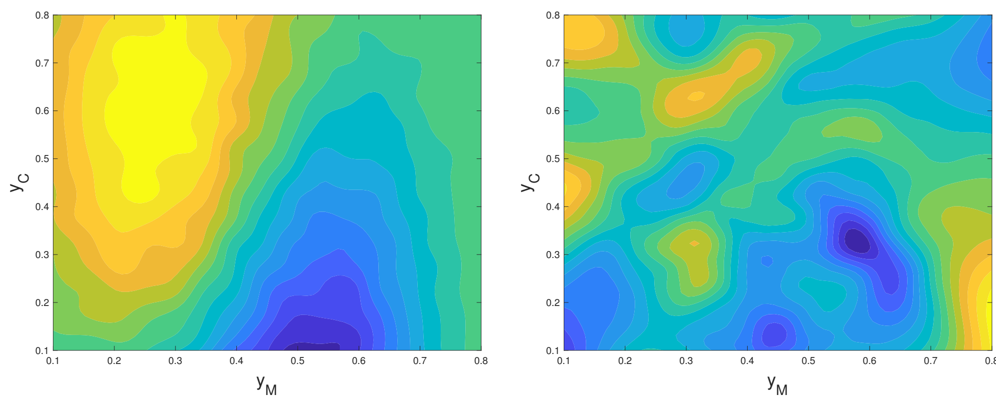


Figure 55: $\Delta\mathcal{F}_t(\mathbf{y})$ vs. Δ in Equilibrium $\mathbf{y}$ from 1979 to 2010 (III)

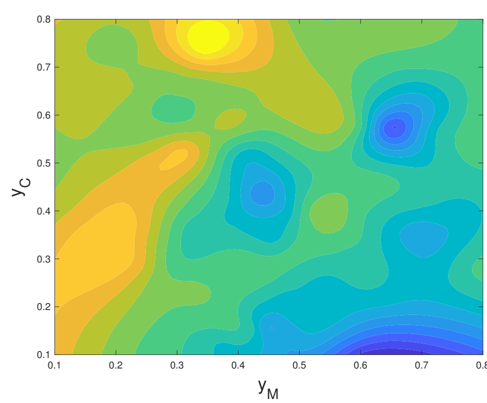


Figure 56: Δ in Equilibrium $\mathbf{y}$ from 1979 to 2010 (Data)